5-2015

Firing Threats: Incentive Effects and Impression Management

Brice Corgnet
Chapman University, corgnet@chapman.edu

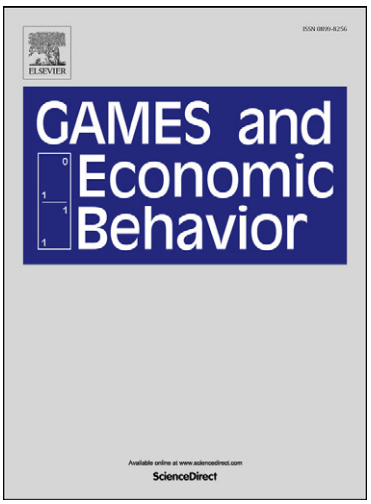
Roberto Hérnan-Gonzalez
Chapman University

Stephen J. Rassenti
Chapman University, rassenti@chapman.edu

Recommended Citation

Corgnet, B., Hernán-González, R., Rassenti, S., 2015. Firing threats: Incentive effects and impression management. *Games and Economic Behavior* 91, 97–113. doi:10.1016/j.geb.2015.02.015

Firing Threats: Incentive Effects and Impression Management



Firing threats: Incentive effects and impression management

Brice Corgnet, Roberto Hernán-González, Stephen Rassenti

PII: S0899-8256(15)00038-X
DOI: http://dx.doi.org/10.1016/j.geb.2015.02.015
Reference: YGAME 2406

To appear in: *Games and Economic Behavior*

Received date: 10 January 2014

Please cite this article in press as: Corgnet, B., et al. Firing threats: Incentive effects and impression management. *Games Econ. Behav.* (2015), http://dx.doi.org/10.1016/j.geb.2015.02.015

Highlights

- We study the effect of firing threats in a real-effort/real-leisure experiment.
- Firing threats sharply increase worker production and decrease on-the-job leisure.
- Workers are inclined to impress their boss in the presence of firing threats.
- Firms endowed with firing threats underperform those using individual incentives.
- After the removal of firing threats, production drops and on-the-job leisure surges.

Firing Threats: Incentive Effects and Impression Management

Brice Corgnet[a], Roberto Hernán-González[b,c,d,*], and Stephen Rassenti[b]

[a] George L. Argyros School of Business and Economics, Chapman University, US
[b] Economic Science Institute, Chapman University, US
[c] Departamento de Teoria e Historia Economica, Universidad de Granada, Spain
[d] Business School, University of Nottingham, UK

Abstract

We study the effect of firing threats in a virtual workplace that reproduces features of existing organizations. We show that organizations in which bosses can fire up to one third of their workforce produce twice as much as organizations for which firing is not possible. Firing threats sharply decrease on-the-job leisure. Nevertheless, organizations endowed with firing threats underperformed those using individual incentives. In the presence of firing threats, employees engage in impression management activities to be seen as hard-working individuals in line with our model. Finally, production levels dropped substantially when the threat of being fired was removed, whereas on-the-job leisure surged.

KEYWORDS**:** Firing threats, incentives, impression management

JEL CODES: C92, D23, D82.

The authors acknowledge financial support from the International Foundation for Research in Experimental Economics, the Argyros School of Business and Economics at Chapman University, the Spanish Ministry of Education [Grant 2012/00103/001], Spanish Plan Nacional I+D MCI [ECO2013-44879-R], 2014-17, and Proyectos de Excelencia de la Junta Andalucía [P12.SEJ.1436], 2014-18. This research was developed when the second author was visiting the Economic Science Institute at Chapman University and working at the University of Granada.

* Corresponding author at University of Nottingham, Business School, Division of Industrial Economics and Finance, Nottingham, NG8 1BB, UK.
E-mail addresses: corgnet@chapman.edu (B. Corgnet), roberto.hernan@nottihgham.ac.uk (R. Hernán-Gonzalez), rassenti@chapman.edu (S. Rassenti).

1. INTRODUCTION

In settings in which employers are unable to provide individual incentives to workers, the threat of being fired becomes an essential feature of an employment contract (Becker and Stigler, (1974), Klein and Leffler, (1981), Shapiro and Stiglitz, (1984), MacLeod and Malcomson, (1989)). At the empirical level, researchers have attempted to assess the effectiveness of firing threats by studying employment protection legislation. For example, Ichino and Riphahn (2005) study absenteeism levels of the workers of an Italian bank before and after a probationary period. The authors report an increase in the level of absenteeism after the probationary period suggesting a negative incentive effect of employment protection. Similarly, Riphahn (2004) shows a negative relationship between employment protection and the level of absenteeism of German workers. Recently, Jacob (2013) reports a 10% decrease in absenteeism levels after principals were allowed more flexibility in firing teachers in the public schools system in Chicago. Interestingly, Jacob (2013) distinguishes incentive and selection effects of the newly implemented policy, and shows the prevalence of selection over incentive effects. These studies provide field evidence that restricting dismissal policies is likely to foster absenteeism which could in turn reduce workers' performance.

In this paper, we propose a different approach by assessing the impact of firing threats in a laboratory environment in which we directly assess the incentive effects of firing threats. In addition to obtaining direct measures of productivity and shirking behavior, the laboratory setting allows us to control for possible confounding factors such as firm size, industry structure, job characteristics, demand shocks, and organizational design.

Our aim is to complement previous research by measuring directly workers' effort, productivity and shirking behavior and assess the magnitude of incentive effects which were found to be modest in the field (Jacob, 2013). To that end, we study firing threats within a virtual workplace that reproduces features of existing organizations such as real effort tasks, real-time monitoring, on-the-job leisure (Internet browsing) and chatting (Corgnet, Hernan-Gonzalez and Rassenti (2013)).[1] We study organizations in which bosses are endowed with a real-time monitoring technology so as to assess the work of their nine employees in each of the five

[1] This experimental platform was built in line with previous research introducing real-effort experiments in the study of labor issues (e.g. Dickinson (1999), Van Dijk, Sonnemans, and van Winden (2001)).

periods of the experiment. In addition, we gave organizational members access to an electronic chat room to exchange messages during the experiment.

We consider three types of incentive structures. In all cases subjects received the same fixed wage at the beginning of each period regardless of pending work productivity. Under the fixed wage treatment employees received no further incentives, while under firing threats they could be fired. Under individual incentives employees could not be fired and each employee was rewarded the entire income generated by his or her individual production on the work task. In the fixed wage and firing threats treatments the boss kept all the income generated by all members of the organization. Under firing threats, the boss was also given the option to dismiss one employee at the end of each of periods 2, 3, and 4.[2] Bosses saved on labor costs after firing employees as they would not have to continue to pay their fixed wages. Subjects who made it to the start of the last period without being fired found themselves with de facto tenure for that final production period.

Our analysis relates to the seminal work of Brown, Falk and Fehr (2004) which studies a repeated principal-agent model à la Fehr, Kirchsteiger and Riedl (1993) in which there is an excess supply of agents. In this setting, principals and agents can sign one-period contracts which specify a fixed payment from the principal to the agent and a desired but non-enforceable level of transfer from the agent to the principal. A crucial difference with the original setting of Fehr et al. (1993) is that the authors allow for reputational concerns and long-term contracts by keeping subjects' identification numbers constant across periods. The authors find that principals and agents were willing to develop long-term relationships which in turn resulted in high levels of transfers. The findings in Brown et al. (2004) are in line with the disciplining version of the efficiency wage hypothesis according to which a combination of high wages and threat of dismissal leads to high levels of effort.

In a recent experimental study, Falk, Huffman and MacLeod (2011) extend the work of Brown et al. (2004) by introducing barriers to dismissals. The authors show that dismissal barriers tend to deter principals from building long-term relationships with agents. This is the case because agents' transfers are significantly reduced when the threat of being dismissed by the principal is eliminated. Note that in Brown et al. (2004) and in Falk et al. (2011) dismissals occur either

[2] Workers could not be rehired by the boss in the rest of the experiment.

because the principal signs a one-period contract with another agent or because the agent quits. Even though this contractual design constitutes a privileged setting for studying relational contracts and dismissal barriers, it cannot isolate the effect of firing threats from the effect of quitting. In the present study, we abstract away from career concerns and labor markets and focus on the impact of firing threats on organizational behavior.

Outside the laboratory, a number of studies using archival data have documented pervasive incentive effects by comparing work performance under hourly wages and piece rates (e.g. Lazear (2000), Shearer (2004)). Also, recent research using data from a large US firm during the 2008 recession suggests that the intensification of firing threats during the economic downturn may explain the increase in productivity and work effort during the period (Lazear, Shaw and Stanton (2013)). In contrast to previous research, we study firing threats in a laboratory environment in which we can directly control for the intensity of firing threats as well as observe managers' firing decisions. In our setting, we assess the effect of different incentive schemes on workers' behavior inside firms. This includes the analysis of work effort, on-the-job leisure and workplace communication in addition to standard measures of work performance such as productivity.

We found that organizations in which bosses were allowed to fire their employees produced twice as much as organizations which only relied on the payment of fixed wages. This was the case even though by the end of the experiment organizations which could fire employees were about 30% smaller than those that could not. Firing threats also decreased Internet usage and chatting activities by 77.7% in those periods in which the threat of being fired was present. Remarkably, firing threats reduced leisure activities and increased production levels for both low- and high- ability workers. Nevertheless, the incentive effect of firing threats was not as compelling as those of individual incentives, as organizations endowed with individual incentives outperformed those endowed with firing threats by 43.3%. Also, the implementation of firing threats required significantly more monitoring effort (12.5% of managers' time) than the implementation of individual incentives (only 2.5% of managers' time) making it a less appealing option for managers. Interestingly, leisure activities were as low in the presence of firing threats as they were under individual incentives. As a result, the difference in workers' production levels between the two treatments was due to a discrepancy in productivity levels rather than to a difference in working time. These findings suggest that in the presence of firing

threats, employees were willing to signal themselves as hard-working individuals who spend long hours at their workstation without browsing the Internet. Social psychologists refer to this process by which people attempt to influence others' perceptions of themselves as *impression management*.[3]

When firing threats disappear during the last production period of the firing threats treatment, workers' production collapsed, and Internet usage surged, to levels which were similar to those of organizations which solely relied on the payment of fixed wages.

2. EXPERIMENTAL DESIGN

2.1. *Virtual Organizations*

We develop a framework in which subjects can undertake a real-effort organizational task while having access to Internet browsing and chatting activities at any point in time during the experiment. Each activity was undertaken in a separate window so that the experimenter had a precise measurement of the time spent on each activity. We consider organizations with ten subjects, nine of which were referred to as *B* subjects (employees) while the remaining subject was referred to as the *C* subject (boss). *C* subjects could monitor *B* subjects' activities in real time. A session consisted of 5 periods of 20 minutes each.[4] The experimental environment is described in detail below.[5]

2.1.1. *The Work Task*

We introduced a particularly long and laborious task so as to ensure that completing the *work task* required a significant level of effort. Subjects were asked to sum up matrices of 36 numbers for 1 hour and 40 minutes.[6] As a result, we expected to identify signs of fatigue and boredom during the experiment. 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 complete tables

[3] According to Newman (2009, p. 184), *impression management* is an "act presenting a favorable public image of oneself so that others will form positive judgments".

[4] Due to a technical issue with the software, one of the sessions in Treatment W lasted only for 4 periods. This problem was solved for the other sessions. We control for this effect in the analysis of the results.

[5] See instructions at http://sites.google.com/site/vofiring/instructions. A video presentation of the software is available at http://sites.google.com/site/vofiring/videos.

[6] Different variations of this task have been used by Bartling, Fehr, Maréchal and Schunk (2009), Dohmen and Falk (2010), and Abeler, Falk, Goette and Huffman (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.

correctly. Each table had 6 rows and 6 columns. The numbers in each table were generated randomly.

Each table completed correctly generated a 40-cent profit while a penalty of 20 cents was subtracted from individual production for each incorrect answer.[7] At the end of each period, and only then, the total amount of money generated by all *10* subjects during the period was displayed in the history panel located at the bottom of their screens.

2.1.2. *Internet Browsing*

The Internet browser was embedded in the software so that the experimenter could keep a record of the switching times between activities as well as the exact amount of time subjects spent on each activity. Subjects were informed that their usage of the Internet was strictly confidential.[8]

The introduction of 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 completed online. Indeed, a large number of people are likely to derive utility from Internet access as they will be able to browse Web pages that best fit their personal interests.[9]

2.1.3. *Chatting Activities*

Subjects 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 a subject 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 his or her screen. As a result, incoming chat could potentially distract subjects completing the *work task*.

2.1.4. *Monitoring Activities*

In all treatments, the *C* subject could monitor the nine *B* subjects' activities at any time during the experiment. Monitoring activities had to be undertaken in a separate window so that the boss

[7] Penalties did not apply when individual production was equal to zero so that individual production could not be negative.
[8] Subjects were expected to follow the norms set by the university regarding the use of Internet on campus.
[9] Two related studies (Eriksson, Poulsen and Villeval (2009), Charness, Masclet and Villeval (2010)) have also introduced on-the-job leisure activities in experimental environments by giving subjects access to magazines.

could not complete his or her own *work task*, *chat* or *browse the Internet* while monitoring his or her employees. In the monitoring screen, the boss could decide whether to monitor all or a subset of the employees at the same time. 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). Note that this technology assumes that bosses are able to monitor all the non-work activities of employees with great precision.[10] One may argue that in a field setting, workers may be able to shirk without being observed by their monitor. However, current technology (e.g., SpectorSoft 360, Virtual Monitoring™, Employee Monitoring) allows for real-time monitoring that closely resembles our experimental design by allowing monitors to scrutinize employees activities and track the time spent on various applications.

B subjects were notified with a message stating "*The C subject is watching you*" jointly with an eye picture whenever they were being watched. At the end of each period, the *C* subject had access to a monitoring summary, including information regarding *B* subjects' production levels and contributions to total production. We acknowledge that in practice, workers may not necessarily be aware that they are being monitored. This lack of monitoring awareness could thus reduce the salience of firing threats potentially affecting their effectiveness.[11]

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. This aimed at representing the pay that workers obtain just for being present at their workstation regardless of their commitment to the *work task*. 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 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.

[10] Using cell phones or reading material unrelated to the experiment was not allowed in the lab so that on-the-job leisure activities were precisely measured by the experimenter.

[11] However, monitoring awareness is also likely to have a positive effect under fixed wages (Mas and Moretti, 2009) and low-powered incentives schemes (Corgnet et al. 2013). The fact that production levels were twice higher under firing threats than under fixed wages is thus unlikely to be accounted for by our methodological choice of making workers aware of monitoring.

2.2. *Treatments*

We conducted three different treatments (see Table 1). In Treatments W (fixed wage) and F (firing threats), *B* subjects were rewarded a fixed payment of 200¢ each period while not receiving incentives based on their performance on the *work task*. The *C* subject received the output produced by all subjects (including him or herself) on the *work task* while not being paid any fixed wage. In Treatment F, the *C* subject could fire one *B* subject at the end of each of periods 2, 3 and 4. The *C* subject kept the fixed pay of dismissed *B* subjects in the following periods.

TABLE 1. Summary of the treatments.

Treatment	Description	Number of sessions (subjects)
Fixed wage (W)	*B* subjects were paid a fixed wage of 200¢ per period. The *C* subject kept the value of all output produced by all *B* subjects in the organization. In addition the *C* subject was paid the value of his/her own production. The *C* subject could monitor *B* subjects' activities but had no possible recourse.	5 (50)[12]
Firing threats (F)	The *C* subject could monitor *B* subjects' activities, and could fire one *B* subject at the end of periods 2, 3 and 4. Payment as in (W) but the *C* subject also kept the fixed pay of dismissed *B* subjects.	6 (60)
Individual incentives (I)	*B* subjects were paid a fixed wage of 200¢ per period and were also rewarded the full value of their production. The *C* subject was paid only the value of his/her own production. The *C* subject could monitor *B* subjects' activities but had no possible recourse.	6 (60)

Dismissed *B* subjects 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 B* and *C* subjects.[13] They were not able to chat with active *B* and *C* subjects, and they could not be rehired.

In Treatment I, *B* subjects received a fixed payment of 200¢ per period as in the previous treatments, in addition to being rewarded on the *work task* according to their individual production. In this treatment, *C* subjects were paid only the value of their own production (they did not collect the value of the output produced by *B* subjects). Treatment I is designed to lead to

[12] One session was cancelled because of insufficient subjects.
[13] As a result, the maximum period earnings of dismissed subjects on the clicking task were equal to 48¢ instead of 240¢ for active *B* and *C* subjects.

the maximum level of production by paying *B* subjects the full product of their effort, that is a payment of 40¢ per correct table. Under individual incentives, no residual output can be allocated to *C* subjects because *B* subjects are already paid the full product of their effort.[14] We use Treatment I as a benchmark to assess the incentive effects of firing threats.[15]

In all treatments, subjects received their individual earnings on the clicking task.

2.3. *Procedures*

Our subject pool consisted of students from a major US University. The experiments took place in May 2011. In total, 170 subjects participated in the experiment, divided into 17 sessions. We conducted five sessions for Treatment W, and six sessions for each of Treatments F and I. Ten students participated in each session. 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. Participants playing the role of a *B* (*C*) subject in Treatments W, F, and I, earned on average \$28.00 (\$55.25), \$27.74 (\$85.20), and \$38.95 (\$37.91), respectively. This includes a \$7.00 show-up fee. Experimental sessions lasted on average two hours and thirty minutes.

3. HYPOTHESES: FIRING THREATS AND IMPRESSION MANAGEMENT

In order to establish predictions regarding production levels and Internet usage across treatments, we rely on standard incentive theory (see Laffont and Martimort (2002) for a review).

[14] Alternatively, we could have artificially doubled the value of the output generated by *B* subjects in Treatment I and pay 40¢ to *C* subjects for each correct table completed by a *B* subject as is the case in Treatments W and F. However, this alternative design would have made it difficult to compare production levels in Treatment I (where stakes would have been doubled) with production levels in Treatments W and F.

[15] We did not consider the treatment combing individual incentives and firing threats as our main purpose was to study the incentive effects of firing threats in comparison to weak (fixed wages) and strong (individual incentives) incentives schemes. In the theoretical literature, firing threats and individual incentives are studied in isolation since the latter is already supposed to lead to an efficient level of effort. Thus, firing threats are seen as a substitute for individual incentives rather than as a complement. Nevertheless, an interesting avenue for future research would be to study the conditions under which firing threats and individual incentives could complement each other in practice. This study would require conducting an experiment where both, firing threats and individual incentives, are present.

We build on a two-period model of an organization composed of multiple workers and a supervisor.[16] In the firing treatment, the first period corresponds to a probationary period at the end of which the supervisor can fire workers.

In our model, workers can dedicate their time to either work or leisure activities. Workers may possess different levels of ability on the *work task*. Regardless of their ability level, workers vary in their level of intrinsic motivation to complete the task (Deci (1971, 1975) and Deci, Koestner and Ryan (1999)). Workers' levels of ability and intrinsic motivation are assumed to be private information. Note that instead of intrinsic motivation we could have invoked altruism (Rotemberg, 1994) and other social preferences (Fehr and Schmidt, 1999) to justify that workers may exert effort in the absence of monetary incentives. However, in the presence of firing threats workers may downplay distributive concerns and increase production to avoid dismissal.[17]

In addition to working on the task and undertaking leisure activities, workers can be present at their workstation without generating production. This may give their supervisor the impression that they prefer the *work task* to Internet browsing. In the field, these nonproductive activities may include shirking activities which are not perfectly monitored by managers such as using cell phones for personal purposes or reading offline material unrelated to work. In that case, the worker may either decide to complete the task with minimal effort (e.g. providing random answers) or choose not to complete the task at all. We refer to this behavior as *impression management*. The supervisor can monitor workers and obtain information regarding their production levels and the time they spend at their workstation. However, the supervisor cannot distinguish between productive and nonproductive work activities.

In this section, we present the hypotheses derived from our theoretical model. The formal version of the model is presented in Appendix B and proofs are available in Appendix 1 online.

[16] This model does not include multiple probationary periods as is the case in our experimental design. The essence of our results would not be affected by considering such case.

[17] Distributive concerns may also be dampened in our model because payoff asymmetry across roles does not depend on one person's decision but on the production of the ten members of the organization. This makes it difficult for subjects to coordinate on an individual level of production that would limit payoff asymmetry. We may also conjecture that aversion to inequality is reduced in a context in which nine of the ten participants in the experiment still earn the exact same amount regardless of subjects' individual production.

Hypothesis 1 (Production levels and Internet usage)

i) We expect workers' production levels to be greater in the treatments with either individual incentives (I) or firing threats (F) than in the fixed wage treatment (W). Also, we expect workers' Internet usage to be lower in Treatments I and F than in Treatment W.

ii) Workers' production levels as well as Internet usage can be either greater or lower in the individual incentives treatment than in the treatment with firing threats.

A necessary condition for the treatment with firing threats to lead to greater production levels and lower Internet usage than individual incentives is for intrinsic motivation to be sufficiently large. The treatment with individual incentives is expected to lead to greater production levels and lower Internet usage than the treatment with firing threats for low to moderate levels of intrinsic motivation.

By distinguishing between productive and non-productive effort, our theoretical framework allows for *impression management* to arise in equilibrium. For example, consider the case of high-ability non-intrinsically motivated workers who decide to mimic the work behavior of low-ability intrinsically motivated workers in order to reduce their likelihood of being fired (see equilibria of types *i* and *iii* in Table O.1 in Appendix 1 online). High-ability workers will be able to achieve the output of low-ability workers in a fraction of the time necessary for low-ability workers to do so. In the remaining time, high-ability workers will stay at their workstation exerting non-productive effort in order to mimic the work behavior of low-ability workers. A consequence of *impression management* is that individual productivity measured as individual production per unit of working time is expected to be lower in the firing treatment compared with Treatments W and I. We state this conjecture as follows.

Hypothesis 2 (Impression management)

We expect to observe impression management in the treatment with firing threats. As a result, we expect productivity levels to be lower with firing threats than in the other two treatments.

Regarding firing decisions, our theoretical framework predicts that supervisors will dismiss workers with the lowest performance levels since they signal low ability as well as a lack of intrinsic motivation.

Hypothesis 3 (Firing decisions)

We expect supervisors to fire the workers with the lowest levels of performance in the organization.

Following our theoretical framework, we also expect supervisors to shy away from monitoring activities in the treatments without firing threats (Treatments W and I) preferring either to work for cash or to browse the Internet. By contrast, in the firing treatment (F), monitoring workers' production and effort levels is valuable to supervisors as it increases the probability of not dismissing intrinsically motivated workers who will continue to work even after firing threats are removed. We summarize this conjecture as follows.

Hypothesis 4 (Monitoring)

We expect the time spent by the supervisor monitoring workers to be greater in the treatment with firing threats than in the other two treatments.

An important implication of our theoretical framework is that the effect of firing threats is expected to be observed for both low- and high- ability workers.

Hypothesis 5 (Firing threats and workers' ability)

We expect the positive effect of firing threats on workers' production levels and its negative effect on Internet usage to be observed across ability levels.

In the last period of Treatment F, firing threats are removed so that workers' levels of production and Internet usage are expected to converge to the levels which are observed in Treatment W.

Hypothesis 6 (Removal of firing threats)

i) *In the last period of the experiment, we expect the production levels of workers not to be significantly different between the fixed wage treatment with firing threats (F) and the fixed wage treatment (W). We expect the production levels of workers to be greater in the treatment with individual incentives (I) than in the other two treatments.*

ii) In the last period of the experiment, we expect Internet usage not to be significantly different between the fixed wage treatment with firing threats (F) and the fixed wage treatment (W). We expect Internet usage of workers to be lower in the treatment with individual incentives (I) than in the other two treatments.

iii) In the last period of the experiment, we expect the time spent by the supervisor monitoring workers to be the same across all treatments.

4. RESULTS

In Sections 4.1 to 4.5, we analyze the first four periods of our experiment which correspond to probationary periods in Treatment F. In Section 4.6, we present the results of the last period when the threat of being fired is removed. In the appendix, we also report our findings when pooling the data for the five periods of the experiment. However, the interpretation of these results should be taken with care as it includes, in Treatment F, periods in which firing threats apply and the last period in which firing threats are removed.

4.1. *Production and Internet Usage*

We define total production on the *work task* as the total monetary amount generated by a subject's answers on the task divided by the reward for each correct answer (40¢). If not specified otherwise, we refer to the production and Internet usage of workers who are currently part of the organization. This implies that in Treatment F, we exclude the subjects who have been fired before the current period. In Section 4.3, we extend our analysis to include fired workers when discussing selection effects.

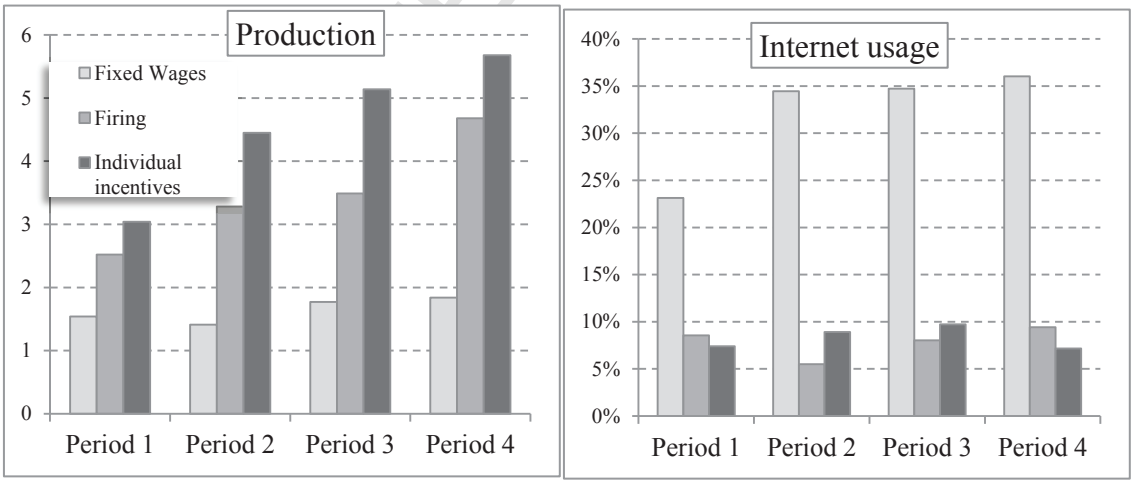


FIGURE 1.— Period-evolution of *B* subjects' average production (left panel) and Internet usage (right panel) across treatments[18]

Figure 1 shows average production (left panel) and Internet usage (right panel) for the first four periods of our three treatments.

[18] Subjects who had been fired before the current period are excluded.

In order to assess any statistical differences in individual production across treatments, we use a series of tests that account for the specific nature of our data. More specifically, we use modifications of standard t-tests and Wilcoxon rank-sum tests to the case of clustered data. We control for clusters at the organization level to take into account a possible correlation in the performance of members of the same organization (see Tables A.4 to A.6).[19] Note that our results continue to hold using standard non-clustered version of the tests.

Production levels for *B* subjects in the first four periods were significantly higher under individual incentives (18.3) than under firing threats (12.8) or fixed wages (6.6) (see Tables A.1 and A.4 in the appendix for a detailed statistical analysis).[20] The fact that individual incentives outperform firing threats is consistent with *Hypothesis 1* for low to moderate levels of intrinsic motivation which is likely to be the case for the dull and repetitive task used in our study. It is interesting to note that in the fixed wage treatment, not only was average individual production strictly greater than zero but the great majority of subjects (82.0%) completed at least one correct table per period. This is consistent with the assumption that workers may be intrinsically motivated to complete the task in the absence of incentives as we account for in our theoretical framework.[21]

In addition, *B* subjects were browsing the Internet significantly more in the fixed wage treatment (31.8% of their time) compared with the firing treatment (8.0%) or the individual incentives treatment (8.3%) (see Tables A.2 and A.5 in the appendix).[22] We identify no significant differences in Internet usage between Treatments I and F.

[19] The p-values of the clustered t-tests can also be obtained by running linear panel regressions with clustered standard errors and a treatment dummy to assess differences in the dependent variable across treatments. The clustered version of the Wilcoxon rank-sum test was performed using Datta and Satten test (2005). Datta and Satten (2005) as well as Galbraith, Daniel and Vissel (2010) provided us with R codes for the test. The codes for the clustered t-test in R were provided by Frank Harrell who implemented the procedure used in Donner, Birkett and Buck (1981).

[20] We do not present treatment comparisons for the case of *C* subjects in the main text because of their limited number per treatment. These results are shown in the appendix.

[21] In an independent survey, we report evidence of the importance of intrinsic motivation in the performance of subjects in our experimental environment. In particular, we show that the answer to the question "How much do you like mathematics?" was significantly correlated with subjects' performance on the *work task* even after controlling for subjects' ability levels (see Appendix 2 online for more details).

[22] Circumstantially, the proportion of their time subjects dedicated to Internet usage under fixed wages (31.8%) was similar to the figures published in the 2005 study by *American Online* and *Salary.com* according to which employees spend about 26.1% of their time on activities unrelated to their work (Malachowski (2005)). Our results show that, even in a laboratory environment usually prone to generating demand effects, subjects were ready to undertake leisure activities for which they were not paid by the experimenter.

Our findings support *Hypothesis 1* according to which firing threats tend to raise production levels and reduce Internet usage compared with the fixed wage treatment.

Despite significant differences in production levels between the firing and the individual incentives treatments, Internet usage was remarkably similar (see Table A.5 in the appendix). As a result, subjects spent as much time on the *work task* screen in the firing treatment (88.3%) as subjects did in the individual incentives treatment (90.0%) while producing significantly less (see Figure 1). In the fixed wage treatment, *B* subjects spent only 47.5% of their time on average on the *work task* screen (see Tables A.3 and A.6 for a detailed statistical analysis of working time differences across treatments and periods).

In other words, not only did subjects produce less in the firing treatment compared with the individual incentives treatment, but they also engaged in significantly more impression management. We assess productivity by computing average individual production for ten minutes of working time. Subjects of type *B* produced on average 1.8, 1.9 and 2.6 tables for ten minutes of working time in the fixed wage, firing and individual incentives treatments, respectively. It is important to note that working time in Treatment W was twice as low as in Treatments F and I implying that productivity in Treatment W could not be directly compared with the other two treatments. This is the case because learning effects on the *work task* are significant as is illustrated by the positive trend in production in both Treatments F and I (see Figure 1).[23] Thus, in order to test *Hypothesis 2* we use a regression analysis in which we compare productivity levels across treatments while controlling for working time. In line with *Hypothesis 2*, we find that productivity levels in Treatment F were significantly lower than in the other two treatments (see Table A.7 in the appendix).

In addition to Internet browsing, chatting activities could divert subjects' attention from the *work task*. Chatting had no strategic function in this experiment except for the firing treatment in which case *C* subjects could use the chat room to define their firing policy and threaten *B* subjects. The categorization of chat messages support the fact that the chat room was largely used as a distraction device (see Appendix 4 online). Interestingly, we observe that, in the treatment with firing threats, *B* subjects excluded the *C* subject from their messages in the

[23] Using a Tobit regression with random effects with individual production as dependent variable and a constant and a trend as regressors, we find that the p-value associated with the trend coefficient was lower than 0.0001 for Treatments F and Treatment I. See also a related study by Corgnet et al. (2013) for evidence of learning effects on a similar task.

majority of cases (54.8%). By contrast, this happened in only a small proportion of the cases in Treatments W (6.9%) and I (9.3%).[24] In the firing treatment, *B* subjects may have excluded the *C* subject from their chat discussions so as to avoid being caught by their boss chatting instead of working.[25] This suggests that firing threats led subjects to feel concerned not only about their production levels but also about their dedication to the *work task*. Employees were then inclined to influence positively the perception of their boss regarding their work dedication. This behavior can be classified as *impression management* (see Newman (2009)). These results relate to the field study of Jacob (2013) that reports a decrease in teachers' absenteeism after firing restrictions had been removed.[26]

Finally, subjects could also obtain earnings from clicking on boxes appearing at the bottom of their screen. No significant differences were observed across treatments regarding the clicking task. Subjects successfully clicked on the box in 94%, 96% and 95% of its appearances in Treatments W, F and I, respectively (see Appendix 3 online).

4.2. *Firing Decisions*

We turn to the analysis of firing decisions of the *C* subjects in Treatment F. We observe that the *B* subjects who were fired in a given period were producing on average significantly less than the other *B* subjects in the organization (see Table 2). In line with *Hypothesis 3*, all the subjects who were fired in periods 2, 3 or 4 were the lowest producers in their respective sessions. In the cases in which several subjects produced the lowest amount on the *work task* (this occurred in four occasions), the subject who was finally fired was the one who was caught browsing the Internet more often during the period.

Also, there is anecdotal evidence that chat conversations may have affected the *C* subject's firing decisions. This was apparent in the first firing decision in Session 1 in which case two subjects produced the same low amount on the task and were both caught on the Internet. Subject B19 was fired after expressing the following opinions publicly in the chat room: "*if C*

[24] We confirm the significance of these findings with a proportion test comparing Treatment F with Treatments W and I (p-value < 0.001). The p-value for the comparison between Treatments W and I is equal to 0.150.

[25] The messages excluding the *C* subjects were mostly general and non-strategic messages (66.5%, 70.0% and 76.5% of messages in Treatments W, F and I).

[26] In the presence of firing threats, teachers may decide to limit evident shirking behaviors by attending school more often while not increasing their work effort. This interpretation seems to be consistent with the fact that the decrease in absenteeism was most pronounced for teachers whose shirking behavior was most noticeable (more than fifteen annual absences). Also, Jacob (2013) reports modest incentive effects which may be due to teachers attempting to impress principals while being reluctant to work harder.

keeps being a rude passive aggressive boss; we can go on strike. If C fires all of us; C makes no income either".

TABLE 2. Summary of firing decisions per period.

	Period 2	Period 3	Period 4	Total
Total [maximum possible] number of fired subjects	5 [6]	5 [6]	4 [6]	14 [18]
Average production of subjects before being fired[27]	0.1	0.9	1.5	6.0
Average production of other *B* subjects	3.2	3.3	3.5	19.8
P-value[28]	0.1972 **[0.031]**	**0.000 [0.000]**	**0.000 [0.000]**	**0.000 [0.000]**

4.3. *Selection Effects*

The selection effect of firing threats follows from the previous finding that worst performers are fired first. This implies that the organizations which are endowed with firing threats retain better workers on average than the organizations assigned to the other two treatments. However, the difference in production levels across treatments does not seem to be solely accounted for by a selection effect. For example, average production was still higher in Treatment F than in Treatment W even after taking into account dismissed workers, who produced nothing in periods 3 and 4 (see row "including fired subjects" in Tables A.1 and A.4).

Additionally, we propose a technique to isolate the selection effect present in Treatment F. To do so, we simulate which workers would have been fired in Treatments W and I had the managers been allowed to dismiss subjects. This amounts to excluding from the analysis the worst performer in period 2 and in period 3 for the remaining periods. Using this procedure, we can compare average production across treatments for organizations which have the same number of producers in all periods.[29] We continue to report a significantly higher average production (all p-values < 0.002), and lower levels of Internet usage (all p-values < 0.005), in the

[27] By multiplying these numbers by 40¢ one obtains the average monetary contribution of those subjects. It is evident that the average monetary contribution is well below the fixed wage of 200¢ received by *B* subjects at the beginning of each period.

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

[29] That is, the 9 *B* subjects in periods 1 and 2, the 8 best performers in period 3, and the 7 best performers in period 4. Given that no workers were fired in one session of Treatment F, we ran 5 (6) different simulations by selecting one session in which we did not exclude any workers in Treatment W (I). P-values do not differ substantially across simulations.

firing treatment compared to the fixed wage treatment. We also find a significantly greater average production in the first four periods in the individual incentives treatment than in the firing treatment (all p-values < 0.087), whereas Internet usage was not significantly different (all p-values > 0.784). In sum, treatment differences in production levels and Internet usage cannot be solely accounted for by selection effects.

4.4. *Monitoring*

The fact that *C* subjects fired *B* subjects according to their relative performance levels suggests that *C* subjects were monitoring *B* subjects to gather information about their production and work dedication. Following *Hypothesis 4* we expect *C* subjects to monitor *B* subjects more intensively in Treatment F than in Treatments W and I (see Table 3). We confirm this conjecture in a Tobit regression with a treatment dummy for firing threats (p-value = 0.031, See Table A.8 in the appendix).[30]

TABLE 3. 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
Fixed wage (W)	5.3%	7.6%	7.0%	2.6%	4.1%
Firing threats (F)	12.2%	12.1%	18.2%	8.7%	9.8%
Individual incentives (I)	2.5%	0.8%	3.4%	4.3%	1.7%

4.5. *The Effect of Firing Threats Across Ability Levels*

We investigate *Hypothesis 5* by analyzing the effect of firing threats on subjects' production across ability levels. To do so, we classify *B* subjects as either high- or low- ability workers depending on whether they completed their first table correctly (we obtain similar results by categorizing subjects according to their performance rank in the first period of the experiment). In an independent study, first table performance was shown to be significantly correlated to subjects' summation skills as measured using an incentivized summation task in the spirit of Dohmen and Falk (2011) (see Table O.4, column 2 in Appendix 2 online). We also rely on previous research showing the positive relationship between first table performance and

[30] When compared separately using Tobit regressions with treatment dummies, the average time spent by *C* subjects monitoring was significantly greater in the firing treatment than in the individual incentives treatment (p-value=0.040). However, we did not find significant differences between the fixed wage treatment and the treatment with individual incentives (p-value=0.233) or between the fixed wage treatment and the firing treatment (p-value=0.201).

subsequent production (see Corgnet et al., 2013 and Appendix 2 online). We classify 65% of the *B* subjects as high-ability workers. This proportion is equal to 58%, 59% and 73% for Treatments W, F and I, respectively. We recognize that this ability measure may have been influenced by treatment effects. We do not find, however, significant differences in the proportion of high-ability *B* subjects across treatments (the p-values for the proportion tests comparing Treatments W and F, W and I, and F and I, are equal to 0.7604, 0.3912, and 0.189, respectively).

Both low- and high- ability *B* subjects produced more in the treatment with firing threats (8.3 and 16.3) than in the fixed wage treatment (3.3 and 8.3) (see Table A.9 in the appendix for statistical analysis). In addition, Internet usage was reduced by 81.0% and 68.0% in Treatment F with respect to Treatment W for low- and high- ability subjects, respectively. In sum, in line with *Hypothesis 5*, the positive effect of firing threats on individual production and its negative impact on Internet usage held across ability levels.

4.6. *Removal of Firing Threats*

In the treatment with firing threats, average production decreased from 4.7 in the fourth period to 3.5 in the last period reaching a level closer to the fixed wage treatment (2.1). Average production in the last period was significantly larger in the individual incentives treatment (6.1) than in the firing treatment. We report no significant differences in last-period average production between the fixed wage treatment and the treatment with firing threats (see Table A.4 in the appendix). As a result, removing firing threats in Treatment F led to a collapse in *B* subjects' production levels. This result is in line with *Hypothesis 6i* as well as with the experimental findings in Falk, Huffman and MacLeod (2011).

In addition, Internet usage increased sharply from an average of 8.0% in the first four periods to 20.9% of total available time in the last period in the firing treatment (p-value < 0.001).[31] Internet usage did not increase significantly in the other two treatments. Similarly, work dedication decreased from an average of 88.3% in the first four periods to 62.0% of total available time in the last period in the firing treatment (p-value < 0.001) (see Figure 2). In the

[31] The reported p-value corresponds to the dummy regressor that takes value one for the last period and value zero for the previous periods in a Tobit panel regression with random effects. The dependent variable is the time spent on Internet by a given subject in a given period and the independent variables are the constant term and the last period dummy. Only the *B* subjects who had not been fired were included in the regression. For Treatments W and I, the p-values of the last period dummy were equal to 0.900 and 0.144, respectively. A similar conclusion can be obtained by using dummies for each period (see Table A.10 in the appendix).

last period, work dedication was significantly lower in the firing treatment (62.0%) and in the fixed wage treatment (41.7%) than in the treatment with individual incentives (88.5%) (see Table A.6 in the appendix). Our findings are consistent with *Hypothesis 6ii*.

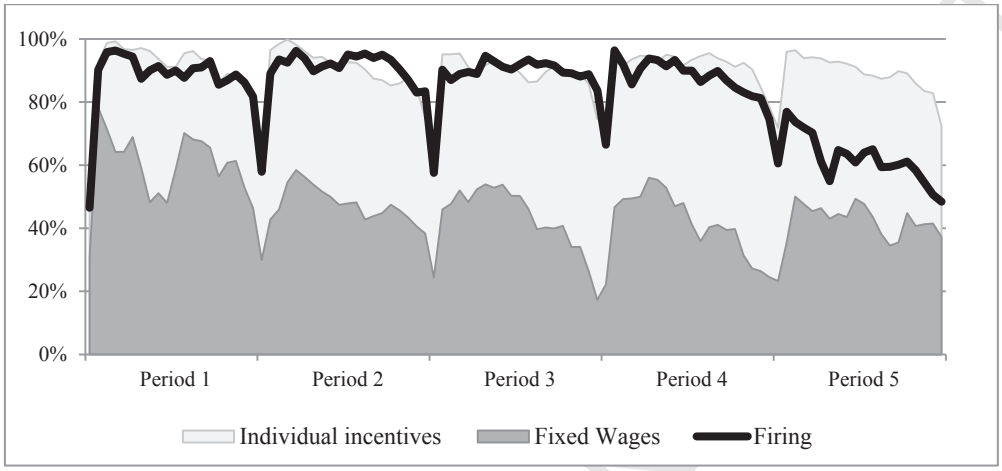


FIGURE 2.— Evolution of the average time (in %) that *B* subjects, who had not (yet) been fired, spent working.

Finally, monitoring time decreased sharply in the firing treatment from an average of 12.2% in the first four periods to 5.2% in the last period. In line with *Hypothesis 6iii*, we find no significant differences in monitoring time between the treatment with firing threats and the other two treatments in the last period (see Table A.8 in the appendix).

5. CONCLUSIONS

In this paper, we investigated the impact of firing threats in a virtual organization characterized by real-effort tasks, access to leisure activities and real-time supervision. We showed that the introduction of firing threats significantly affected organizational behavior. In particular, production was more than twice as high in the presence of firing threats as in their absence while on-the-job leisure activities (Internet browsing and chatting) were almost eradicated. These results show that even though firing threats were limited to a maximum of one-third of the labor force (three out of nine workers), they positively affected the work effort of all employees.

Nevertheless, organizations endowed with firing threats produced significantly less than organizations in which individual incentives were used. Interestingly, organizations endowed with firing threats did not differ from those using individual incentives in terms of leisure activities. These results suggest that employees facing firing threats were willing to appear to their bosses as hard-working individuals spending most of their time on the *work task*. According

to our theoretical framework, a consequence of such signaling behaviors is the reduced productivity of workers in the firing treatment compared with the fixed wage and individual incentives treatments. We were able to confirm this conjecture. The time spent on the *work task* was the same under firing threats and individual incentives while production levels were about 25% lower under firing threats compared to individual incentives. These findings suggest that under firing threats, employees were willing to signal themselves as hard-working individuals who spend long hours at their workstation without browsing the Internet. This implies that firing threats are likely to promote *impression management* activities. As a result, employers may fail to dismiss unmotivated workers leading to a collapse in firm performance when firing threats are removed. This result suggests that previous research reporting a decrease in absenteeism levels after the removal of employment protection (e.g. Riphahn 2004, Jacob, 2013) may be partly explained by the fact that workers try to impress their supervisors by showing up for work more often while not being more productive at work.

Our study uncovers the magnitude of incentive effects of firing threats, complementing earlier field research on employment protection. Previous field research mostly focused on indirect measures of performance such as absenteeism and reported only modest incentive effects.

In this paper, we focused on the incentives effects of firing threats after controlling for screening effects. Nonetheless, the screening effect of firing threats may be substantial as low performers may be reluctant to accept jobs which involve high risks of termination (Jacob, 2013). An interesting avenue for future research would be to study the magnitude of screening effects in a similar controlled environment.

6. REFERENCES

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

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

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

Brown, M., A. Falk, and E. Fehr (2004): "Relational contracts and the nature of market interactions," *Econometrica* 72 (3), 747780.

Charness, G., D. Masclet, and M. C. Villeval (2010): "Competitive Preferences and Status as an Incentive: Experimental Evidence," IZA Discussion Paper 5034, Bonn, and GATE WP 10-16. halshs-00497974.

Corgnet, B, R. Hernán-González, and S. Rassenti (2013): "Peer Pressure and Moral Hazard in Teams: Experimental Evidence," *Review of Behavioral Economics,* forthcoming.

Datta, S., and G. Satten (2005): "Rank-sum Tests for Clustered Data," *Journal of the American Statistical Association*, 100, 908-915.

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., 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(3), 627-668.

Dickinson, D. (1999): "An Experimental Examination of Labor Supply and Work Intensities," *Journal of Labor Economics*, 17, 638-670.

Dohmen, T., and A. Falk (2010): "You Get What You Pay For: Incentives and Selection in the Education System," *The 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.

Donner. A., N. Birkett, and C. Buck (1981): "Randomization by Cluster: Sample Size Requirements and Analysis," *American Journal of Epidemiology*, 114, 906-914.

Eriksson, T., A. Poulsen, and M. C. Villeval (2009): "Feedback and Incentives: Experimental Evidence," *Labour Economics*, 16 (6), 679-688.

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 B. MacLeod (2011): "Institutions and contract enforcement," Discussion Paper No. 361, GESY.

Falk, A., and A. Ichino (2006): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24(1), 39-58.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993): "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics*, 108(2), 437-59.

Fehr, E., and K. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817–868.

Galbraith, S., J. Daniel, and B. Vissel (2010): "A Study of Clustered Data and Approaches to its Analysis," *Journal of Neuroscience*, 30, 10601-10608.

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.

Ichino, A., and R. Riphahn (2005): "The Effect of Employment Protection on Worker Effort: Absenteeism During and After Probation," *Journal of the European Economic Association*, 3(1), 120-143.

Jacob, B. (2013): "The Effect of Employment Protection on Teacher Effort," *Journal of Labor Economics*, 31(4), 727-761.

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.

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

Lazear, E. (2000): "Performance Pay and Productivity," *American Economic Review*, 90(5), 1346-1361.

Lazear, E., K. Shaw, and C. Stanton (2013): "Making Do with Less: Working Harder During Recessions," National Bureau of Economic Research Working Paper Series, 2013, No. 19328.

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

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

Mas, A., and E. Moretti (2009): "Peers at Work," *American Economic Review*, 99(1), 112-145.

Newman, D. M. (2009): "Sociology: Exploring the Architecture of Everyday Life," Thousand Oaks, CA: Pine Forge Press.

Riphahn, R. (2004): "Employment Protection and Effort Among German Employees," *Economics Letters*, 85(3), 353-357.

Rotemberg, J. (1994): "Human Relations in the Workplace." *Journal of Political Economy* 102, 684-718.

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

Shearer, B. (2004): "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," *Review of Economic Studies*, 71(2), 513-534.

Van Dijk, F., J. Sonnemans, and F. van Winden (2001): "Incentive Systems in a Real Effort Experiment," *European Economic Review*, 45, 187-214.

7. APPENDICES

APPENDIX A: ADDITIONAL TABLES AND REGRESSION ANALYSES

TABLE A.1 Average (median) [standard deviation] individual production across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage	1.54	1.41	1.77	1.84	6.57	2.07	8.56
		(1.00)	(1.00)	(1.00)	(1.00)	(4.00)	(0.75)	(6.00)
		[1.75]	[1.61]	[2.39]	[2.45]	[6.52]	[3.09]	[8.71]
	Individual Incentives	3.04	4.45	5.14	5.68	18.31	6.10	24.41
		(2.50)	(4.00)	(5.50)	(5.00)	(15.25)	(6.00)	(21.25)
		[2.76]	[3.92]	[3.87]	[3.96]	[13.21]	[4.21]	[16.93]
	Firing Threats							
	All 10 subjects (including fired subjects)	2.52	3.28	3.17	3.81	12.78	2.58	15.36
		(2.25)	(2.50)	(2.75)	(3.50)	(12.25)	(1.00)	(14.25)
		[2.56]	[3.18]	[2.72]	[3.97]	[11.52]	[3.58]	[14.38]
	Excluding fired subjects[#]	-	-	3.49	4.68	12.87	3.49	15.65
				(3.50)	(4.00)	(12.58)	(2.25)	(15.25)
				[2.64]	[3.91]	[11.48]	[3.76]	[14.27]
C Subjects only	Fixed Wage	2.00	2.80	3.30	5.60	13.70	6.88	19.20
		(2.00)	(1.50)	(3.00)	(5.00)	(9.50)	(7.25)	(16.00)
		[1.58]	[2.75]	[2.54]	[3.58]	[9.78]	[3.75]	[14.02]
	Individual Incentives	3.67	4.50	5.67	6.00	19.833	7.17	27.00
		(3.00)	(4.00)	(3.50)	(5.25)	(15.00)	(5.50)	(20.50)
		[4.19]	[4.60]	[6.27]	[5.19]	[19.34]	[6.68]	[25.80]
	Firing Threats	2.08	2.83	4.50	3.83	13.25	3.33	16.58
		(2.50)	(2.00)	(3.50)	(3.50)	(9.75)	(3.00)	(12.75)
		[1.32]	[3.27]	[3.62]	[3.24]	[9.89]	[1.86]	[11.67]

[#]In this case, total (subtotal) production was computed as the total production in periods 1 to 5 (1 to 4) divided by the number of periods each worker was part of the company and multiplied by 5 (4). Note that this procedure is equivalent to calculating the sum of workers' production across periods for the other treatments. We used the same procedure to calculate total production in Treatment W, given that in one session subjects could participate in the first four periods only, due to an error in the software.

TABLE A.2 Average (median) [standard deviation] percentage of time subjects spent on Internet across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage	23.14 (7.93) [29.52]	34.45 (24.90) [32.73]	34.72 (20.99) [32.54]	36.04 (28.58) [32.57]	31.80 (29.57) [24.67]	33.69 (29.59) [31.35]	31.06 (29.96) [22.78]
	Individual Incentives	7.40 (2.30) [12.31]	8.91 (1.70) [17.90]	9.73 (1.48) [19.91]	7.16 (0.94) [17.42]	8.30 (2.45) [14.70]	10.21 (1.32) [21.32]	8.68 (2.78) [14.99]
	Firing Threats							
	All 10 subjects (including fired subjects)	8.55 (2.58) [15.63]	5.49 (2.01) [10.75]	16.54 (2.01) [31.51]	26.19 (1.87) [38.53]	14.19 (2.57) [19.54]	41.40 (18.46) [43.38]	19.63 (8.46) [22.22]
	Excluding fired subjects	-	-	8.02 (1.90) [17.22]	9.42 (1.44) [16.58]	7.97 (2.28) [13.19]	20.89 (2.90) [29.88]	9.93 (3.05) [14.06]
C Subjects only	Fixed Wage	4.17 (4.65) [1.81]	5.53 (1.52) [7.49]	4.71 (3.03) [4.78]	5.57 (3.06) [5.73]	4.70 (4.30) [3.70]	6.76 (2.14) [10.11]	4.89 (4.93) [3.29]
	Individual Incentives	1.66 (1.79) [0.49]	3.40 (1.61) [3.78]	3.50 (1.48) [4.99]	3.09 (1.76) [3.53]	2.91 (1.87) [1.86]	3.80 (1.58) [6.12]	3.09 (1.78) [2.46]
	Firing Threats	2.40 (2.20) [0.72]	2.42 (2.50) [0.84]	5.00 (2.28) [5.69]	1.32 (1.19) [0.61]	2.78 (2.14) [1.64]	2.14 (1.36) [1.82]	2.66 (2.16) [1.61]

TABLE A.3 Average (median) [standard deviation] percentage of time subjects spent working across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage	59.61 (65.99) [29.96]	46.69 (43.29) [33.56]	42.51 (45.48) [35.06]	41.19 (39.63) [35.16]	47.49 (50.41) [27.60]	41.69 (33.86) [40.02]	44.94 (43.51) [27.56]
	Individual Incentives	90.93 (97.27) [13.42]	89.51 (98.04) [18.46]	88.37 (97.77) [20.50]	91.21 (98.37) [17.62]	90.00 (96.66) [15.29]	88.53 (98.05) [21.60]	89.71 (96.22) [15.47]
	Firing Threats							
	All 10 subjects (including fired subjects)	87.87 (95.69) [18.25]	89.90 (96.19) [14.49]	80.37 (96.73) [32.98]	70.86 (91.85) [38.94]	82.25 (93.73) [21.38]	45.91 (39.32) [42.79]	74.98 (81.51) [23.38]
	Excluding fired subjects	-	-	88.57 (97.13) [21.44]	86.96 (98.33) [20.88]	88.30 (95.82) [16.61]	61.98 (73.60) [38.30]	83.46 (90.02) [18.97]
C Subjects only	Fixed Wage	76.68 (85.30) [23.89]	74.06 (84.13) [21.71]	72.54 (89.50) [28.05]	76.33 (82.73) [22.90]	75.39 (89.33) [20.97]	87.88 (87.99) [12.27]	74.96 (86.82) [20.65]
	Individual Incentives	97.39 (98.02) [1.73]	92.68 (97.59) [8.30]	91.38 (95.99) [13.68]	95.07 (96.29) [4.69]	94.13 (97.00) [5.23]	89.59 (93.89) [11.58]	93.22 (96.97) [6.22]
	Firing Threats	81.62 (83.54) [16.30]	69.59 (71.41) [25.99]	82.01 (80.04) [9.60]	82.50 (85.81) [16.74]	78.93 (78.56) [12.64]	89.87 (93.19) [10.40]	81.12 (81.79) [11.34]

TABLE A.4 P-values for clustered t-tests (clustered Wilcoxon rank-sum tests) assessing differences in period production across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage vs. Individual Incentives	**0.012 (0.042)**	**0.000 (0.007)**	**0.000 (0.007)**	**0.000 (0.006)**	**0.000 (0.007)**	**0.000 (0.009)**	**0.000 (0.007)**
	Fixed Wage vs. Firing Threats							
	All 10 subjects (including fired subjects)	**0.033 (0.066)**	**0.000 (0.013)**	**0.038 (0.044)**	**0.004 (0.032)**	**0.002 (0.037)**	0.587 (0.744)	**0.005 (0.041)**
	Excluding fired subjects	**-**	**-**	**0.017 (0.027)**	**0.000 (0.009)**	**0.001 (0.035)**	0.135 (0.165)	**0.006 0.039**
	Individual Incentives vs. Firing Threats							
	All 10 subjects (including fired subjects)	0.335 (0.360)	0.118 (0.133)	**0.018 (0.054)**	**0.040 (0.032)**	**0.051 (0.054)**	**0.000 (0.004)**	**0.009 (0.020)**
	Excluding fired subjects	-	-	**0.015 (0.058)**	**0.106 (0.077)**	**0.055 (0.061)**	**0.000 (0.012)**	**0.012 (0.025)**
C Subjects only	Fixed Wage vs. Individual Incentives	0.928 (0.999)	0.986 (0.782)	0.536 (0.926)	0.418 (0.464)	0.517 (0.931)	0.155 (0.133)	0.748 (0.855)
	Fixed Wage vs. Firing Threats	0.400 (0.645)	0.470 (0.647)	0.427 (0.782)	0.884 (0.999)	0.941 (0.999)	0.634 (0.713)	0.543 (0.792)
	Individual Incentives vs. Firing Threats	0.411 (0.809)	0.488 (0.518)	0.703 (0.999)	0.409 (0.520)	0.481 (0.748)	0.226 (0.518)	0.398 (0.688)

TABLE A.5 P-values for clustered t-tests (clustered Wilcoxon rank-sum tests) assessing differences in Internet usage across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage vs. Individual Incentives	**0.012 (0.042)**	**0.000 (0.007)**	**0.000 (0.007)**	**0.000 (0.006)**	**0.000 (0.004)**	**0.000 (0.014)**	**0.000 (0.005)**
	Fixed Wage vs. Firing Threats							
	All 10 subjects (including fired subjects)	**0.005 (0.029)**	**0.000 (0.013)**	**0.006 (0.029)**	(0.191) **(0.062)**	**0.000 (0.006)**	0.370 (0.624)	**0.000 (0.009)**
	Excluding fired subjects	**-**	**-**	**0.000 (0.016)**	**0.000 (0.018)**	**0.000 (0.006)**	0.229 (0.156)	**0.000 (0.009)**
	Individual Incentives vs. Firing Threats							
	All 10 subjects (including fired subjects)	0.753 (0.695)	0.322 (0.773)	0.189 (0.337)	**0.001 (0.040)**	**0.079** (0.248)	**0.000 (0.011)**	0.267 (0.411)
	Excluding fired subjects	-	-	0.690 (0.661)	0.672 (0.156)	0.929 (0.717)	0.188 (0.178)	0.775 (0.646)
C Subjects only	Fixed Wage vs. Individual Incentives	**0.095 (0.082)**	0.408 (0.931)	0.930 (0.931)	0.172 (0.329)	0.360 (0.537)	0.430 (0.609)	0.229 (0.429)
	Fixed Wage vs. Firing Threats	**0.009 (0.034)**	0.585 (0.931)	0.690 (0.999)	0.428 (0.662)	0.328 (0.537)	0.624 (0.690)	0.345 (0.429)
	Individual Incentives vs. Firing Threats	**0.067** (0.132)	0.559 (0.818)	0.637 (0.937)	0.485 (0.690)	0.910 (0.937)	0.548 (0.999)	0.745 (0.937)

TABLE A.6 P-values for clustered t-tests (clustered Wilcoxon rank-sum tests) assessing differences in working time ─or leisure time─ across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1 to 4	Period 5	TOTAL
B Subjects only	Fixed Wage vs. Individual Incentives	**0.000 (0.003)**	**0.000 (0.004)**	**0.000 (0.003)**	**0.000 (0.004)**	**0.000 (0.002)**	**0.000 (0.011)**	**0.000 (0.002)**
	Fixed Wage vs. Firing Threats							
	All 10 subjects (including fired subjects)	**0.000 (0.006)**	**0.000 (0.003)**	**0.000 (0.005)**	**(0.000) (0.010)**	**0. 000 (0.003)**	0.745 (0.680)	**0.000 (0.007)**
	Excluding fired subjects	-	-	**0.000 (0.004)**	**0.000 (0.006)**	**0.000 (0.003)**	0.161 (0.129)	**0.000 (0.006)**
	Individual Incentives vs. Firing Threats							
	All 10 subjects (including fired subjects)	0.493 (0.156)	0.924 (0.201)	0.135 (0.248)	**0.001 (0.061)**	**0.062** (0.145)	**0.000 (0.010)**	**0.001 (0.020)**
	Excluding fired subjects	-	-	0.971 0.574	0.491 (0.278)	0.713 (0.452)	**0.007** (0.138)	0.295 (0.167)
C Subjects only	Fixed Wage vs. Individual Incentives	0.125 **(0.004)**	0.130 (0.126)	0.223 (0.126)	0.142 (0.126)	0.131 **(0.082)**	0.832 (0.762)	0.138 **(0.082)**
	Fixed Wage vs. Firing Threats	0.707 (0.662)	0.763 (0.931)	0.505 (0.999)	0.631 (0.931)	0.739 (0.931)	0.798 (0.762)	0.668 (0.931)
	Individual Incentives vs. Firing Threats	**0.064 (0.065)**	**0.083 (0.093)**	0.203 (0.240)	0.129 **(0.093)**	**0.041 (0.093)**	0.965 (0.699)	**0.067 (0.065)**

TABLE A.7 Tobit regression with random effects for individual productivity (periods 1 to 4) across treatments.[32]

	Treatment F vs. I	Treatment F vs. W
Intercept	-0.016	-0.081**
Treatment F	-0.067***	-0.050**
Working time (in minutes)	0.014***	0.017***
Number of observations and Log likelihood (L)	n = 480 (63 left censored) L = -24.777, [Prob > χ^2] < 0.001	n = 440 (76 left censored) L = -25.176, [Prob > χ^2] < 0.001

Treatment F is a dummy variable that takes value 1 for Treatment F and 0 otherwise
*p -value<.10, ** p-value<.05, and *** p-value<.01.

TABLE A.8 Tobit regression with random effects for monitoring time (in seconds).

	Periods 1-4[33]	Period 5
Intercept	11.553	-10.867
Treatment F	115.707**	64.667
Number of observations and Log likelihood (L)	n = 68 (18 left censored) L = -668.2 [Prob > χ^2] = 0.046	n = 16 (6 left censored) L = -132.0 [Prob > χ^2] = 0.175

Treatment F is a dummy variable that takes value 1 for Treatment F and 0 otherwise
*p -value<.10, ** p-value<.05, and *** p-value<.01.

[32] These results are robust to alternative specifications including non-linear specifications of working time.
[33] These results are robust to introducing period dummies.

TABLE A.9 P-values for clustered t-tests (clustered Wilcoxon rank-sum tests) assessing differences in individual production and Internet usage across treatments and ability levels for periods 1 to 4.

	Individual production per period		Internet usage per period	
	Low-ability	High-ability	Low-ability	High-ability
Fixed Wage vs. Individual Incentives	**0.012** **(0.051)**	**0.002** **(0.016)**	**0.000** **(0.002)**	**0.006** **(0.011)**
Fixed Wage vs. Firing Threats				
All 10 subjects (including fired subjects)	**0.068** (0.376)	**0.014** **(0.021)**	**0.015** **(0.022)**	**0.011** **(0.016)**
Excluding fired subjects	**0.065** (0.296)	**0.012** **(0.020)**	**0.000** **(0.004)**	**0.008** **(0.013)**
Individual Incentives vs. Firing Threats				
All 10 subjects (including fired subjects)	0.200 (0.129)	0.320 (0.284)	**0.070** (0.166)	0.876 (0.616)
Excluding fired subjects	0.205 (0.178)	0.334 (0.315)	0.882 (0.487)	0.984 (0.712)

TABLE A.10 Tobit regression with random effects for Internet usage per period for *B* subjects only.

	Fixed Wages (W)	*Firing (F)*	*Individual incentives (I)*
Intercept	280.475***	102.570***	88.823***
Period 2	136.353**	-36.681	18.136
Period 3	136.119**	-6.200	27.966
Period 4	156.514***	15.515	-2.899
Period 5	148.247**	117.186***	33.660
Number of observations and Log likelihood (L)	$n = 216$ (5 right censored) L = -9913.297 [Prob > χ^2] < 0.001	$n = 238$ (0 right censored) L = -3206.362 [Prob > χ^2] < 0.001	$n = 270$ (0 right censored) L = -3624.998 [Prob > χ^2] = 0.498

*p -value<.10, ** p-value<.05, and *** p-value<.01.

APPENDIX B: CONCEPTUAL FRAMEWORK

We build on a two-period model of an organization composed of *n* workers and a supervisor.

Workers

In each period, worker $i \in N = \{1, \dots, n\}$ dedicates his or her time to attend either work ($e_i \geq 0$) or leisure activities ($l_i \geq 0$) where $e_i + l_i = \varphi$ and $\varphi > 0$ is the total amount of time available to workers. We allow for workers who are present at their workstation to dedicate their time either to productive ($e_i^P \geq 0$) or nonproductive activities ($e_i^{\sim P} \geq 0$) where $e_i = e_i^P + e_i^{\sim P}$. Only productive effort generates production on the *work task* (q_i). However, workers may still decide to attend their workstation without generating production ($e_i^{\sim P} > 0$) so as to give their supervisor the impression that they prefer the *work task* to Internet browsing. In that case, the worker may either decide to complete the task with minimal effort (e.g. providing random answers) or choose not to complete the task at all. We refer to this behavior as *impression managemen*t.

Definition. Worker *i* is involved in *impression management* whenever $e_i^{\sim P} > 0$.

We assume that each worker *i* possesses a level of ability on the *work task* denoted by $\alpha_i \in \{\alpha_L, \alpha_H\}$, with $\alpha_L < \alpha_H$, which determines the marginal product of the productive effort as $q_i = \alpha_i e_i^P$. We assume that ability levels are workers' private information. We denote by $N_j \subset N$ the set of workers endowed with ability α_j, with $j \in \{L, H\}$. We denote by $n_j > 0$ the number of workers endowed with ability α_j so that the total number of workers is defined as $n = n_L + n_H$.

We consider that the cost of productive effort $C(e_i^P)$ for worker *i* is such that $C'(\cdot) > 0$ and $C''(\cdot) > 0$. We assume that the nonproductive effort entails no costs save the opportunity costs of not producing for cash or enjoying the leisure activity. Also, we denote by $v(l_i)$ the utility that worker *i* derived from the leisure activity, where $v'(\cdot) > 0$ and $v''(\cdot) < 0$. In order to keep the focus of our analysis on workers' heterogeneity in abilities, we consider that workers have the same cost of effort and the same utility of leisure. We denote by ω the fixed wage received by each worker at the beginning of a period.

Importantly, we consider that regardless of their ability level workers can either be intrinsically motivated to complete the *work task* or not. We assume that a worker's intrinsic motivation is private information. We denote by $n_{j,R}$ ($n_{j,\sim R}$) the number of (not) intrinsically motivated

workers of ability α_j, with $j \in \{L, H\}$. Intrinsically motivated workers derive direct utility from working (Deci (1971, 1975) and Deci, Koestner and Ryan (1999)). We denote such utility as $\vartheta(e_i^P)$, with $\vartheta(e_i^P) \coloneqq 0$ for non-intrinsically motivated workers, and with $\vartheta(e_i^P) > 0$ for $e_i^P > 0$, $\vartheta(0) = 0$, $\vartheta'(\cdot) > 0$ and $\vartheta''(\cdot) < 0$ for intrinsically motivated workers.

Supervisor

Each organization is monitored by a supervisor $s \notin N$. The supervisor can monitor workers and obtain information regarding their production levels (q_i) and their dedication to the *work task* (e_i). However, the supervisor cannot distinguish between productive and nonproductive work activities. In addition to monitoring, supervisors can dedicate their time either to work or leisure activities so that: $m_s + e_s + l_s = \varphi$ where m_s is the time dedicated to monitoring activities. For simplicity, we consider two levels of monitoring intensity ($m_s \in \{0, \bar{m}\}$).[34] For $m_s = \bar{m}$, the supervisor can observe individual production and effort levels while in the absence of monitoring ($m_s = 0$), the supervisor can only observe the total production of the organization, that is $\sum_{i=1}^{n} q_i + q_s$, where q_s is the supervisor's production. The cost of monitoring activities is specified as follows: $C_m(0) = 0$ and $C_m(\bar{m}) = \bar{c}$ where $\bar{c} > 0$. In all treatments, the supervisor's own effort is rewarded at its marginal product.

Now we describe subjects' utility functions for each treatment:

Individual Incentives Treatment (Treatment I)

In the individual incentives treatment, workers' utility function is described as follows:

$$U_i \coloneqq \sum_{t=1}^{2} \left[\omega + \alpha_i e_i^{P(t)} + v\left(l_i^{(t)}\right) - C\left(e_i^{P(t)}\right) + \vartheta\left(e_i^{P(t)}\right)\right]$$

and supervisor's utility function is as follows:

$$U_s \coloneqq \sum_{t=1}^{2} \left[\alpha_s e_s^{P(t)} + v\left(l_s^{(t)}\right) - C\left(e_s^{P(t)}\right) + \vartheta\left(e_s^{P(t)}\right) - C_m\left(m_s^{(t)}\right) - n\omega\right]$$

Fixed Wage Treatment (Treatment W)

Similarly, in the fixed wage treatment, workers' utility function is described as follows:

[34] We could consider an intermediate level of monitoring for which the supervisor only observes individual production. The nature of our results would not be affected by considering such extension.

$$U_i \coloneqq \sum_{t=1}^{2}\left[\omega + v\left(l_i^{(t)}\right) - C\left(e_i^{P(t)}\right) + \vartheta\left(e_i^{P(t)}\right)\right]$$

and supervisor's utility function is as follows:

$$U_s \coloneqq \sum_{t=1}^{2}\left[\alpha_s e_s^{P(t)} + v\left(l_s^{(t)}\right) - C\left(e_s^{P(t)}\right) + \vartheta\left(e_s^{P(t)}\right) - C_m\left(m_s^{(t)}\right) + \sum_{i\epsilon N} q_i^{(t)} - n\omega\right]$$

Firing Treatment (Treatment F)

In the firing treatment, the first period corresponds to a probationary period at the end of which, and after having paid workers' wages, the supervisor can fire n_f workers where $n_f \epsilon \{0,1,\dots,\bar{n}_f\}$ and $\bar{n}_f \leq n$. We denote N_F the set of fired workers. In the second period, which can be seen as a tenure period, the workers who were not fired at the end of the first period will receive the same fixed wage $\omega > 0$ as in the first period. Fired workers will not receive any fixed wage and will not be able to produce anymore. Workers will not have the opportunity to work in another organization in the second period. This cost of being fired can be seen as temporary unemployment. In that setting, the utility function for worker *i* can be described as follows:

$$U_i \coloneqq \omega + v\left(l_i^{(1)}\right) - C\left(e_i^{P(1)}\right) + \vartheta\left(e_i^{P(1)}\right)$$
$$+(1-\pi_i)\left\{\omega + v\left(l_i^{(2)}\right) - C\left(e_i^{P(2)}\right) + \vartheta\left(e_i^{P(2)}\right)\right\} + \pi_i v\left(l_i^{(2)}\right)$$

where π_i is the probability for worker *i* to be fired at the end of the first period. Supervisor's utility function is as follows:

$$U_s \coloneqq \sum_{t=1}^{2}\left[\alpha_s e_s^{P(t)} + v\left(l_s^{(t)}\right) - C\left(e_s^{P(t)}\right) + \vartheta\left(e_s^{P(t)}\right) - C_m\left(m_s^{(t)}\right)\right]$$
$$+\sum_{j\epsilon N} q_j^{(1)} - n\omega + \sum_{j\epsilon N\setminus N_F} q_j^{(2)} - (n - n_f)\omega$$

We derive our conjectures using the following specification of the model: $C(e_i^P) = \frac{(e_i^P)^2}{2}$, $v(l_i) = \beta l_i$ and for intrinsically motivated subjects we assume that $\vartheta(e_i^P) = \gamma e_i^P$, where $\beta > 0$ and $\gamma > 0$. We provide details of our derivations in Appendix 1 online.