2013

Peer Pressure and Moral Hazard in Teams: Experimental Evidence

Brice Corgnet
Chapman University, corgnet@chapman.edu

Roberto Hérnan-Gonzalez

Stephen Rassenti
Chapman University, rassenti@chapman.edu

Peer Pressure and Moral Hazard in Teams: Experimental Evidence

Comments
Working Paper 13-01

Peer Pressure and Moral Hazard in Teams: Experimental Evidence[1]

Brice Corgnet
Argyros School of Business and Economics, Chapman University

Roberto Hernan-Gonzalez
Universidad de Granada

Stephen Rassenti
Economic Science Institute, Chapman University

Abstract: *Holmström (1982) established that free riding behaviors are pervasive whenever people are paid according to aggregate measures of output such as team incentives. However, team incentives have been found to be particularly effective both in the lab and in the field. In this paper we show, in line with Holmström (1982), that shirking behaviors in teams are indeed pervasive. Production levels were significantly lower under team incentives than under individual incentives while the time dedicated to on-the-job leisure activities (Internet usage) was significantly larger under team incentives than under individual incentives. Subsequently, we find that a very weak form of peer monitoring (anonymous and without physical proximity, verbal threats or face to face interactions) allowed organizations using team incentives to perform as well as those using individual incentives. This provides strong evidence for the conjecture of Kandel and Lazear (1992) that peer pressure may resolve the moral hazard in teams problem.*

KEYWORDS**:** Incentives, free-riding, monitoring, peer pressure, organization theory
JEL CODES: C92, D23, M52

[1] An earlier version of this paper circulated as “Real Effort, Real Leisure and Real-time Supervision: Incentives and Peer Pressure in Virtual Organizations” Economic Science Institute, Chapman University Working Paper 11-05.

An organization can secure the efforts necessary to its existence, then, either by the objective inducements it provides or by changing states of mind. . . . We shall call the process of offering objective incentives "the method of incentives"; and the processes of changing subjective attitudes "the method of persuasion." —Barnard (1938, p. 142)

1. INTRODUCTION

Team Incentives in the Theory of Organizations

As a point of departure for the analysis of organizations and the development of an economic theory of the firm, theorists have put forward the pervasiveness of free-riding behaviors in teams in which it is difficult to observe and verify the contribution of each partner (Alchian and Demsetz (1972), Holmström (1982)). Indeed, workers paid according to an aggregate measure of performance such as team output are likely to exert less effort than if they were paid according to their individual performance. A central feature of successful organizations consists of overcoming free riding by designing effective monitoring schemes (Alchian and Demsetz (1972)) or using budget-breaking devices aimed at threatening potential free riders (Holmström (1982)).

At the empirical level, the evidence of free riding behavior in teams has been limited (e.g. Encinosa, Gaynor and Rebitzer (2007), Leibowitz and Tollison (1980)). Instead, team incentives have been found to be particularly effective both in laboratory experiments (Dohmen and Falk (2011), van Dijk, Sonnemans and van Winden (2001)) and in field studies (Dumaine (1990, 1994), Hamilton, Nickerson and Owan (2003), Hansen (1997), Ichniowski et al. (1996), Ichniowski, Shaw and Prennushi (1997), Kruse (1992), Manz and Sims (1993)). For example, Hamilton, Nickerson and Owan (2003) show that equal sharing of production bonuses within teams seems to stimulate cooperation, information sharing, monitoring and even mutual training, generating a productivity increase (relative to individual incentives) despite the expected free-rider problem. In a recent paper, Babcock et al. (2012) show that team incentives can outperform individual incentives in fostering students' attendance to the university gym Club. The authors acknowledge the crucial role of social pressure in explaining their results.

The empirical difficulty to identify free-riding behaviors in teams is likely due to the lack of control over crucial aspects of work teams that act as confounding factors such as peer monitoring, interpersonal relations or implicit incentives (e.g. firing threats).

In this study, we are able to compare team and individual incentives while controlling for team-specific features that may interfere in the empirical assessment of team incentives. We study large teams (of ten members each) with the aim of recreating a realistic organizational setting. Our experimental environment differs from previous experimental settings as it introduces a long and real-effort *work task* as well as real-time access to leisure activities (Internet browsing). We found the introduction of real-leisure alternative activities to be pertinent as subjects were indeed willing to undertake on-the-job leisure activities for which they were not paid by the experimenter. In particular, subjects spent 11.9% of their time browsing the Internet when they were paid according to individual incentives.

Importantly, we report that production levels were on average 32.8% lower under team incentives than under individual incentives. This result was driven by extensive shirking behaviors in the team incentives treatment in which subjects spent on average 28.5% of their time browsing the Internet. These results are consistent with incentives theory (see Holmström (1979), and Laffont and Martimort (2002) for a review) as they confirm the sound premise that performance is increased by the use of high-powered incentives schemes. As we show in the online appendix (Section A part IV), the introduction of Internet browsing as an on-the-job leisure activity is a crucial feature of our experimental environment as incentives effects mostly vanish in its absence.

As a second step of our analysis, we introduced a real-time monitoring technology in our virtual organizations so as to assess whether the poor performance of team incentives could be mitigated by peer pressure.

Supervision and Peer-monitoring in the Theory of Organizations

Supervision is an important aspect of the theory of the firm that was mentioned by preeminent scholars as one of the raison d'être of organizations (Barzel (1982), Chandler (1992), Jensen and Meckling (1976)). Alchian and Demsetz (1972) put forward the need for centralized supervision in a context of asymmetric information between managers and their subordinates in a team context. They stress that peer monitoring is not an efficient mechanism because the agents would tend to shy away from monitoring activities. However, other theories view peer monitoring as a highly-effective mechanism (Carpenter et al. (2009), Kandel and Lazear (1992)). Kandel and Lazear stress the role of shame arising when workers produce less than the group average as an important mechanism in understanding the effectiveness of peer pressure. Carpenter, Bowles and Gintis (2006) emphasize the role of negative reciprocity as a behavioral mechanism leading contributors to voluntary incur private costs to punish free riders. Evidence of such behaviors has been found in public good experiments (Fehr and Gächter (2000), Sefton, Shupp and Walker (2007)) as well as in contests between groups (Abbink et al. (2010)). Grosse, Putterman and Rockenbach (2011) stress the popularity of peer monitoring devices in a modified version of the public good game in which subjects could vote on whether to use a central monitor or rely on a decentralized monitoring system (peer monitoring). Under peer monitoring, each subject decided how much to invest in the monitoring technology which precision determined the allocation of team profits. In particular, the proportion of the team profits which was allocated according to individual contributions increased in the precision of the monitoring technology. The authors found that subjects mostly relied on peer monitoring as a disciplining device challenging the idea of Alchian and Demsetz that a central monitor is needed to avoid free riding behaviors in the provision of monitoring.

Peer effects have been reported in a series of field experiments. For example, Sacerdote (2001) and Zimmerman (2003) report peer effects on students' grades among college roommates. Falk and Ichino (2006) found that students who worked for fixed wages to stuff envelopes performed significantly better when working in pairs than when working alone. Mas and Moretti (2009) studied the case of supermarket cashiers and found positive peer effects on the number of items scanned by cashiers. The authors considered workers' visual contact and frequency of interactions as measures of peer pressure. In a related field work, Bandiera, Barankay and Rasul (2005) found that mutual monitoring reduced fruit pickers' productivity when they were paid according to relative performance. The authors interpret this result as evidence of workers being aware of the negative effect of achieving high levels of production on their co-workers' pay. In the field studies described previously, peer pressure was approximated by a variety of observable measures such as visual contact, physical proximity or frequency of interactions. In this paper, we bring anonymous real-time supervision in a controlled laboratory environment so as to enable the experimenter to measure peer pressure with precision. In particular, we are able to record the amount of time subjects spent watching others, the activities which were completed by the subjects who were being watched, as well as discern the identity of the subjects who were watching and being watched by others. Furthermore, our anonymous supervision mechanism allows us to isolate the effects of possible cofounds that may appear in a face to face interaction such as fear of retaliation.

Our peer monitoring technology was such that each team member could monitor peers' activities at any point in time during the experiment. As a result, subjects could shape their monitoring strategy by deciding upon which subjects to monitor and when to do so.[2] Monitors were informed in real-time about the activities undertaken by supervisees and could therefore identify whether they were browsing the Internet or producing for the organization. In the peer pressure monitoring treatment, subjects were notified on their screen whenever they were being watched by others. This feature induced social pressure which is defined by Mas and Moretti (2009) as a case in which workers experience disutility if they are observed behaving selfishly by their peers. In that respect, our monitoring technology was more intrusive than the mere release of feedback about relative performance introduced in recent experimental works (Azmat and Iriberri (2010), Blanes i Vidal and Nossol (2011), Eriksson, Poulsen and Villeval (2009), Kuhnen and Tymula (2009)).

Our environment offers a unique opportunity to provide a detailed analysis of peer monitoring activities. In the peer pressure treatment, we report that a large proportion of subjects (88.3%) decided to monitor others. However, subjects dedicated only a small proportion of their time (4.4%) to monitoring activities, compared with the proportion of their time subjects spent working (82.5%) or browsing the Internet (13.1%). Yet, all subjects were being watched for an average of 22.4% of their time. Team members shared the monitoring burden and maintained peer pressure during the whole duration of the experiment. Another important characteristic of the peer monitoring strategy implemented by organizational members was its unpredictability.

We find evidence of strong peer pressure effects when comparing organizations endowed with peer monitoring and team incentives with organizations relying on team incentives alone. Production was 47.1% higher and Internet usage was 54.1% lower under peer pressure. In contrast to public good games with monetary punishments (Carpenter (2007a, 2007b), Fehr and Gächter (2000)), both effort and efficiency were increased by the introduction of peer monitoring. This was the case because subjects spent little time watching others as they shared the monitoring burden to limit the cost of monitoring.

[2] This endogenous aspect of our monitoring technology can be linked to search experiments in which subjects decide whether to observe or not their relative performance (Burks et al. (2010), Falk, Huffman and Sunde (2006)).

Peer monitoring combined with team incentives led to levels of performance and Internet usage that were remarkably similar to individual incentives, despite the absence in our design of punishments devices, communication technologies or physical proximity among subjects.

These findings confirm the theoretical conjecture of Kandel and Lazear (1992) that peer pressure may be an effective solution to the moral hazard in teams problem identified by Holmström (1982). In addition, we were able to answer the question of “How is peer pressure generated? (Kandel and Lazear, 1992, p.805)”. To do so, we conducted experiments in which organizational members could watch each other’s activities without being noticed by their peers. We show that, in contrast to visible peer monitoring, the invisible monitoring technology did not reduce shirking. These results indicate that effective peer monitoring crucially hinges on social pressure.

2. EXPERIMENTAL DESIGN AND HYPOTHESES

2.1. *Virtual Organization*

We develop a framework in which subjects could undertake a real-effort organizational task, have access to Internet, and monitor other subjects’ behavior in real-time.[3]

2.1.1. *The Work Task*

We introduce a particularly long and laborious task so as to reduce as much as possible the role of intrinsic motivation in our environment. Indeed, subjects may like certain tasks and derive direct utility from undertaking the activity. By using a long, repetitive and effortful task we ensure that individual performance is mostly driven by effort considerations. We do so because our main objective is to test standard predictions of incentives theory while abstracting from confounding factors such as intrinsic motivation. The duration of our task as well as its intricacy were considerably higher than in previous real-effort experiments that have reported the use of summation tasks (Eriksson, Poulsen and Villeval (2009), Niederle and Vesterlund (2007), Bartling, Fehr, Maréchal and Schunk (2009) and Dohmen and Falk (2011)). In an early and unmatched contribution to the real-effort literature, Dickinson (1999) designed a four-day experiment in which subjects had to undertake a two-hour typing task. Another long real-effort task was used by Falk and Ichino (2006) in which participants were asked to complete a four-hour mailing task. In a field experiment by Gneezy and List (2006), subjects were asked to enter data (book references) into a computer database for six hours. In our experiment, subjects were

[3] A video presentation of the software is available at http://sites.google.com/site/vopeerpressure/home/videos.

asked to sum up tables of 36 numbers for 1 hour and 40 minutes (see Figure O.1 in the online appendix).

Each table completed correctly generated a 40-cent profit while a penalty of 20 cents was subtracted from individual production for each incorrect answer. After each subject completed a table, the accumulated individual production was updated so that subjects knew whether their answer was correct or not. At the end of each period, and only then, participants were informed about the total amount of money generated by all 10 participants' *work task* during the period.

2.1.2. *Internet Browsing*

At any point during the experiment, participants could switch from the *work task* to the leisure activity that consisted of browsing the Internet. Each activity was undertaken separately, in a different screen. To switch from one activity to another subjects simply had to click on the corresponding option of the drop-down menu at the bottom-right of their screens (see Figure O.2 in the online appendix). In that respect, Internet browsing introduced temptation in the spirit of recent self-control experiments providing on-the-job distraction activities such as watching a humorous video (Bucciol, Houser and Piovesan, 2011). The use of Internet as a tempting alternative in the study of self-control problems has been considered by Houser et al. (2010).

Internet browsing and the *work task* were undertaken on different screens so that subjects could not complete tables while being on the Internet. 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.

The introduction of Internet browsing in our virtual organizations is motivated by the widespread use of Internet in the workplace. According to a 2005 study by *American Online* and *Salary.com*, employees spend about 26% of their time on activities unrelated to their work (Malachowski (2005)). Almost half of this time actually corresponds to Internet usage. In addition, a study by *Nielsen/Net Ratings* report that people spend more than twice as much time online at the office as they do at home (Farrell (2000)). Gordon (2000) argues that Internet usage in the workplace may damage employees' productivity (see also Young (2005, 2006)). Also, the use of Internet represents a suitable alternative to the *work task*, as it is widespread among university students and provides a wide range of activities (Jones et al. (2009).

The consideration of leisure-related issues in the experimental literature was first introduced in the analysis of labor supply by Dickinson (1999). The objective of the author was to assess both income and substitution effects using laboratory experiments. Participants had to undertake a two-hour typing task on four different days (the first day was used for training). In one of the two treatments (the combined experiment), subjects could leave the laboratory whenever they had achieved a certain output level. This aimed at capturing off-the-job leisure activities. Falk and Huffman (2007) also introduced the possibility for subjects to quit the experiment when analyzing minimum wages and workfare in the laboratory. However, it is difficult to interpret the heterogeneity in quitting behaviors given the lack of control over subjects' activities outside the laboratory. Our experimental design embeds on-the-job leisure activities into the work environment that allows the experimenter to measure the exact amount of time each subject spent on leisure and work activities, respectively. Two related studies (Charness, Masclet and Villeval (2010), Eriksson, Poulsen and Villeval (2009)) have also introduced on-the-job leisure activities in experimental environments by giving subjects access to magazines.

2.1.3. *The Low Effort Clicking Task*

In addition to the previously mentioned activities, each subject could click on a yellow box moving slowly from left to right at the bottom of their screen. This *clicking task* aimed at representing the pay that workers obtain just for being present at their workstation regardless of their commitment to the *work task*. These payments can be seen as a fixed wage. The introduction of the *clicking task* allowed subjects to collect a constant flow of earnings with low effort but without actually working dilligently at the high effort *work task*. Each time subjects clicked on a yellow box they earned *5 cents*, no matter how they chose to spend their time otherwise (see Figure O.3 in the online appendix). The box appeared at the bottom of a subject's screen every 25 seconds whether the subject was currently *working on the work task* 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 yellow boxes that appeared on the screen during the experiment.

2.1.4. *Real-time Monitoring*

In the monitoring treatments, subjects were able to monitor others' activities in real time. Our objective was to design an environment that allows for the emergence of peer effects which were defined by Charness and Kuhn (2011) as "…a situation where workers work, side by side, for the same firm but do not interact in any way (except that they observe each others' work activity.)". We allowed subjects to monitor their peers' activities at any time during the experiment by selecting the *Watch* option in the drop-down menu. In that respect, our monitoring technology offered a unique opportunity to assess the effect of peer pressure over time and examine the conjecture that peer effects may fade away as time passes (Falk and Ichino (2006)).

Monitoring activities had to be undertaken in a separate screen so that subjects could not participate in the *work task* or the leisure activity while monitoring others, though they could continue with the *clicking task*. As a result, monitoring imposed an opportunity cost on watchers that was different in nature from the monetary cost of punishments in public good games (Fehr and Gächter (2000)). In the monitoring screen, subjects could decide whether to monitor only a subset or all the other subjects at the same time. Alternative monitoring technologies may be considered in which the monitor can only monitor a subset of subjects at the same time.[4] The information was displayed in a table, where each column showed information regarding the activities completed by a given subject. Monitors were informed about the activities undertaken by each subject (Internet, Work Task or Watch), the current production as well as their contribution to the *work task* (in % terms) and whenever a subject summed a column or a row, before providing a final answer for the *work task* (see Figure O.4 in the online appendix).

In the peer pressure treatment, subjects were notified with a message stating the experiment *ID* of the watcher jointly with an eye picture whenever they were being watched (see Figure O.5 in the online appendix). We also conducted a treatment in which subjects were not notified when they were being watched by others (invisible monitoring) so as to isolate the role of social pressure in the peer monitoring technology. Note that social pressure, though minimal, is not totally eliminated in the invisible monitoring treatment since workers may still feel that they are watched by their peers even if they are not notified about it.

The monitoring technology used in the present paper allows for precise control over the supervision activities which is difficult to obtain in the field (Bandiera, Barankay and Rasul (2005), Falk and Ichino (2006), Mas and Moretti (2009)). For example, we can measure the exact amount of time subjects were being watched by others as well as the amount of time they spent watching others. It is also possible to identify the watchers as well as the subjects who were being watched. The experimenter has also access to the information that was displayed on the watchers' screens at any given time. Importantly, our anonymous supervision mechanism

[4] These more general monitoring technologies are currently under study.

allows us to isolate the effects of possible cofounds that may appear in a face to face interaction such as the fear of retaliation.

Another distinctive feature of our monitoring technology is that subjects could freely decide upon their monitoring strategy. Subjects could choose who to monitor and when to do so. This feature of the supervision technology will allow us to study subjects' monitoring strategies.

Table 1. Summary of the treatments.

Treatment	Description	Number of sessions (subjects)
Individual incentives (**I**)	Subjects were rewarded on the *work task* according to their individual production.	7 (66)
Team incentives (**T**)	Subjects were rewarded on the *work task* by obtaining 10% of the total production of the 10 group members in each session.	6 (60)
Peer pressure (**TP**)	Subjects were rewarded according to team incentives. Subjects had access to a monitoring technology that notified them whenever they were being watched.	6 (60)
Invisible monitoring (**TPN**)	Subjects were rewarded according to team incentives. Subjects had access to a monitoring technology for which they were not notified when they were being watched.	6 (60)

2.2. *Treatments*

We ran four different treatments (see Table 1). In the baseline treatment, subjects were rewarded on the *work task* according to their individual production (treatment *I*). In the second treatment (treatment *T*), the total production of the 10 subjects participating in the experiment was equally distributed among them. Our third experiment was the peer pressure treatment (treatment *TP*) which was equivalent to treatment *T* except that all ten subjects could monitor their peers using the technology described in the previous section. Treatment *TPN* was similar to treatment *TP* except that organizational members were not notified on their screen when they were being watched by another subject.

In all treatments subjects could individually obtain the full rewards ($2.40 per period) of the *clicking task* with minimal vigilance and effort. The instructions for each treatment are available online.[5]

We conducted a series of additional treatments as robustness checks (see Part IV of the online appendix). In particular, we ran two treatments which were similar to treatments *I* and *T* except

[5] http://sites.google.com/site/vopeerpressure/home/instructions. Instructions for treatment *TPN* were the same as for treatment *TP* except for slide 36 which was removed.

for the fact that Internet browsing was not made available to subjects. We also assessed the robustness of our results by controlling for demographic information which was collected in a follow-up study.

2.3. Conceptual Framework

We build our conceptual framework on the moral-hazard in teams problem introduced by Holmström (1982) and on its extension to the presence of peer monitoring which was proposed by Kandel and Lazear (1992). We consider *N* workers producing a total output $f(e_1, e_2, ..., e_N)$ which depends on each worker's effort e_i, where $i \in \{1, ..., N\}$. Each worker i decides to allocate her time to the following activities: work effort $(e_i \geq 0)$, leisure $(l_i \geq 0)$ or peer monitoring $(m_i \geq 0)$. We normalize these variables to one so that $e_i + l_i + m_i = 1$. Following Holmström (1982) and Kandel and Lazear (1992), we consider the case of homogeneous workers and assume the same utility function for all workers. In our case, this implies the same cost of effort, $C(.)$, and the same utility for Internet usage, $\eta(.)$, for all workers. It also implies the same peer pressure function $P(.)$ for all workers. The utility function of worker *i* can be expressed as follows: $U_i := s_i f(e_i; e_j, ..., e_N) - C(e_i) + \eta(l_i) - P(e_i, e_0, m_{-i})$, where $C(e_i)$ stands for the cost of effort function with $C' > 0$ and $C'' > 0$, s_i is the share of group production assigned to worker *i* and m_{-i} is the vector of peer monitoring activities for workers $j \neq i$. Under team incentives (treatments *T*, *TP* and *TPN*), $s_i = \frac{1}{N}$ while under individual incentives (treatment *I*) workers are paid according to their actual contribution to the group outcome. If we assume that $f(.)$ is separable in workers' effort and in particular if we assume that $f(e_i; e_j, ..., e_N) = \sum_{k=1}^{N} e_k$ then $s_i = \frac{e_i}{\sum_{k=1}^{N} e_k}$ under individual incentives.[6] In order to provide an illustration of the peer pressure function we refer to the work of Andreoni and Bernheim (2009) on social image. In particular, we consider that a worker *i* will obtain utility (suffer a utility loss) from being watched by at least

[6] Kandel and Lazear (1992) assume nonseparability in effort so as to justify the existence of partnerships and eliminate the possibility of self-employment. In this paper, we do not aim at justifying the existence of partnerships and simply assume separability of the utility function in effort so as to match our experimental design more closely.

one coworker ($\exists\, j \neq i : m_j > 0$) whenever she is producing more (less) than the benchmark level of effort e_0. In the parlance of Andreoni and Bernheim (2009), our peer pressure function implies that workers care about their social image. Workers feel pride if they produce more than the benchmark and feel shame if they fall short of the benchmark contribution. In order to assess the interaction between audience effects and peer monitoring, we add to the authors' discussion a distinction between visible and invisible audiences. We aim at considering both the case in which workers are aware of others' scrutiny (treatment *TP*) and the case in which they are not (treatment *TPN*). The degree to which the audience is visible is denoted by the parameter v which is equal to v_{TP} and v_{TPN} in treatments *TP* and *TPN*, respectively. We assume that $v_{TP} > v_{TPN}$. That is, we assume that a person is more affected by social image concerns when the audience is visible than when it is not. We specify the peer pressure function as follows:

$-P(e_i, e_0, m_j) := v_\mu \chi (e_i - e_0)(N-1) m_j$, where $i \neq j$, $\mu \in \{TP, TPN\}$ and $\chi \geq 0$ captures the extent to which worker *i* cares about her social image. As in Kandel and Lazear (1992), we take into account that workers are ex-ante identical and will choose the same level of monitoring activity m_j. We derive our main predictions by using the following specification of the workers' utility function.[7]

$U_i := s_{i,R} \sum_{k=1}^{N} e_{k,R} - \alpha \frac{e^2_{i,R}}{2} - \beta \frac{(1 - l_{i,R})^2}{2} + v_R \chi (e_{i,R} - e_0)(N-1) m_{j,R}$ where $\alpha > 0$, $\beta > 0$, and $R \in \{I, T, TP, TPN\}$ represents the experimental treatment in which worker *i* was involved. By definition, $m_{j,T} = m_{j,I} = 0$ since peer monitoring is not available in the individual incentives and in the team incentives treatments. We obtain the following equilibrium values for work effort for each treatment.[8]

$$e^*_I = \frac{1}{\alpha + \beta} \qquad e^*_T = \frac{1}{N(\alpha + \beta)} \qquad e_\mu^{\ *} = \frac{\frac{1}{N} + \left[v_\mu \chi (N-1) - \beta \right] m_\mu^{\ *}}{\alpha + \beta} \tag{1}$$

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

[8] We derive these calculations in the online appendix (Part II). Note that we have to impose $\alpha + \beta > 1$ for e^*_I not to be strictly greater than one.

Our conceptual framework which models peer pressure as the result of audience effects is closely related to the concept of *social facilitation* according to which a person's performance on a given task is likely to be affected by the presence of others (Zajonc (1965). In particular, Zajonc stresses that the presence of others affects performance positively for simple and well learnt tasks. We summarize our findings in the following hypotheses.

Hypothesis 1 (Individual incentives versus team incentives)

Production is expected to be greater and Internet usage is expected to be lower under individual incentives than under team incentives.

Regarding the comparison of the team incentives and the peer pressure treatments, standard incentives theory would predict no differences both in terms of production and Internet usage. In contrast with the *work task*, subjects had no monetary incentives to monitor others. Peer monitoring was a time consuming activity during which monitors had either to sacrifice *work task* earnings or leisure time. As a result, in the absence of social image concerns $(\chi = 0)$, we expect subjects to shy away from monitoring activities. We would then expect treatments *T* and *TP* to lead to the same levels of production and Internet usage. However, in the presence of concerns for social image $(\chi > 0)$, it follows from our theoretical framework that work contribution in the peer monitoring treatments can be larger than in the team incentives treatment without peer monitoring. Indeed, as long as concerns for social image are sufficiently large $(v_\mu \chi > \frac{\beta}{N-1})$ peer monitoring ($m_\mu^* > 0$ in equation (1)) will lead to an increase in work effort (e_μ^*). This prediction is stated in the following hypothesis.

Hypothesis 2 (Peer monitoring)

i) Production is expected to be greater and Internet usage is expected to be lower in the peer monitoring treatments, $\mu \in \{TP, TPN\}$, *than in the team incentives treatment without peer monitoring as long as workers are sufficiently concerned with their social image* ($v_\mu \chi > \frac{\beta}{N-1}$).

In that case, workers are expected to dedicate part of their time to peer monitoring activities so as to foster the effort of the other workers.

ii) Production is expected to be greater in the peer monitoring treatments than in the individual incentives treatment for very high levels of social image concerns.

We should also recognize that one might expect peer monitoring activities to backfire generating distrust among workers. Recent research has emphasized this negative aspect of monitoring and put forward that trusting employees can lead to higher levels of effort than intensive supervision (Dickinson and Villeval (2008), Falk and Kosfeld (2006), Fehr, Klein and Schmidt (2007a, 2007b), Frey (1993)). We do not consider crowding-out of effort as our primary hypothesis because the disciplining effect of supervision has been found to be dominant in the absence of interpersonal relationships among workers as is the case in our experimental design (Dickinson and Villeval (2008), Frey (1993)). In addition, crowding-out effects are likely to be stronger in a principal-agent relationship or in any situation in which the monitor has some authority on the supervisee's work. In our design, we consider a multi-agent monitoring structure in which there is no principal and no hierarchy among subjects.

Finally, our theoretical setting assumes that the impact of social image concerns is diminished under invisible monitoring compared with the peer pressure treatment since $v_{TP} > v_{TPN}$. As a result, we expect the peer pressure treatment to outperform invisible monitoring.

Hypothesis 3 (Peer pressure and invisible monitoring)

Production is expected to be lower and Internet usage is expected to be greater in the invisible monitoring treatment than in the peer pressure treatment.

Interestingly, the invisible monitoring treatment will also help us assess whether any effect of the peer monitoring technology can be accounted for by the access to continuous feedback on others' production levels. In particular, if the effect of peer monitoring on workers' production levels is driven by the access to feedback rather than to social pressure we should expect invisible monitoring to perform as well as the peer pressure treatment. Also, the comparison of the invisible monitoring treatment and the team incentives treatment without monitoring will inform us on the importance of feedback in the effectiveness of peer monitoring technologies (see Nikiforakis (2010) for the study of feedback in public good games with monetary punishments).

2.4. *Procedures*

Our subject pool consisted of students from a major American university with a diverse population. Participants were recruited by emails from a pool of more than 2,000 students who had signed up to participate in experiments. Emails were sent on a random basis to a subset of the pool of students. The experiments took place in December 2010 and February 2011. In total, 246 subjects participated in the experiment, divided in 25 sessions. We ran seven sessions for treatment *I*, and six sessions for each of treatments *T*, *TP*, and *TPN*. Ten students participated in each session, except for two sessions of 8 students that corresponded to treatment *I*. The experiment was computerized using the *Virtual Organizations* software proprietarily developed for the authors. All of the interaction was anonymous.

The instructions were displayed on subjects' computer screens.[9] Subjects had exactly 20 minutes to read the instructions. A 20-minute timer was shown on the laboratory screen. Three minutes before the end of the instructions period, a monitor entered into the room announcing the time remaining and handing out a printed copy of the summary of the instructions. None of the participants asked for extra time to read the instructions. At the end of the 20-minute instruction round, the experimenter closed the instructions file from the server, and subjects typed their names to start the experiment. The interaction between the experimenter and the participants 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 are computed as the sum of the earnings in the *5* periods. Participants in treatments *I*, *T*, *TP*, and *TPN* earned on average $27.25, $24.45, $27.10, and $24.95 respectively. This includes a $7.00 show-up fee. Experimental sessions lasted on average two hours and fifteen minutes.

3. RESULTS

We start by comparing individual and team incentives (Hypothesis 1). We study peer monitoring treatments (Hypotheses 2 and 3) subsequently. Additional results are provided in part III of the online appendix while robustness checks are conducted in part IV of the online appendix.

3.1. *Team Incentives Versus Individual Incentives*

3.1.1. *Individual Production*

We define production as the monetary amount generated by a subject's answers on the *work task* divided by the reward for each correct answer (40¢). It can be interpreted as the total number of correct tables completed by a given subject discounted by the number of incorrect answers. In both treatments, period production steadily increased except for the third period as is

[9] Subjects were told that all screens displayed the same set of instructions.

illustrated in Figure 1 (see Table III.1 in the online appendix for regression analyses). The summation task considered in the current experiment was significantly longer and more complex than in the works of Niederle and Vesterlund (2007) or Erikson, Poulsen and Villeval (2009) in which no learning effects were reported. In our task, subjects could develop strategies to sum up the 36 numbers in the table at a faster speed. For example, subjects could decide not to compute the partial sum of rows and columns and compute only the final sum of all the numbers in the table. As a result, it is not surprising to observe a positive trend in production levels which can be seen as evidence of a learning effect which has been identified in long arithmetic tasks (Charness and Campbell (1988)).

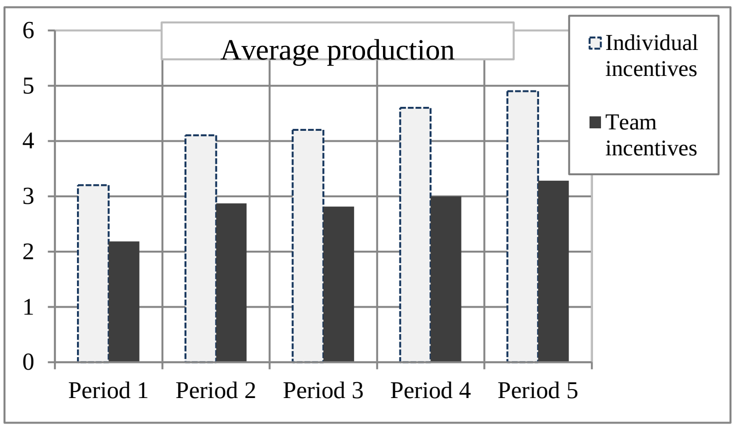

Figure 1. Average production per period across treatments.

Average individual production per period was equal to 4.21 tables under individual incentives compared with 2.83 tables under team incentives. This corresponds to a 48.8% production gap between the individual incentives treatment and the team incentives treatment. Interestingly, this gap was observed for each of the five periods.

The comparison of individual production across treatments stresses that organizations using individual incentives significantly outperformed those using team incentives regardless of the test we used (see Table A.1 in the appendix).[10] This finding also holds when comparing

[10] We use standard t-tests and Wilcoxon rank-sum tests as well as modifications of these tests to the case of clustered data. The clustered version of the Wilcoxon rank-sum test was performed using Datta and Satten test (2005) while the clustered version of the t-test followed Donner, Birkett and Buck (1981). We aim at controlling for the fact that individual production in a given session may be affected by group production. This correction is especially relevant for the treatment with team incentives in which case the contributions of other group members, displayed on a subject's screen at the end of each period, may affect an individual's motivation. This may have led subjects to free ride whenever they observed an increase in group production as is the case in standard public good games (see Ledyard (1995) for a survey, and see Table III.2 in the online appendix for further analyses).

individual production across treatments for each of the five periods separately (see Table A.2 in the appendix). We summarize our findings as follows:

RESULT 1 (*Work task* production: Individual versus team incentives).
Total production was significantly greater in the individual incentives treatment than in the team incentives treatment. This result also holds when analyzing each period separately.

Result 1 is not surprising in the light of incentive theory (Hypothesis 1) but constitutes an essential step in the empirical analysis of incentives given the limited evidence of free riding behaviors in teams. To our knowledge, this is the first time this result is established in a controlled environment. A related analysis was conducted by Nalbantian and Schotter (1997) in an abstract experimental setting in which the authors compared different types of group incentives programs ranging from revenue sharing (team incentives) to target-based and team-tournament incentives. The authors report that group incentives based on competition among teams outperform other group incentives schemes.[11] In a recent field experiment, Bandiera, Barankay and Rasul (2012) provide a comparison of different group incentives schemes. In particular, they study the effect of group incentives schemes on workers' selection of their team partners. They show that high-powered group incentives lead workers to select their teammates on the basis of ability instead of friendship.

3.1.2. *Internet Usage*

We report a positive trend in Internet usage in both treatments. Internet usage increases significantly from period 2 onwards under individual incentives while it is not until period 3 that Internet usage took off under individual incentives (see Table III.3 in the online appendix).

Under team incentives subjects spent on average 28.5% of their time browsing the Internet while this percentage was only equal to 11.9% under individual incentives.[12]

[11] In this paper, we do not study different types of group incentives schemes. Rather, we focus on team incentives (revenue sharing) schemes. Notice that Nalbantian and Schotter (1997) did not compare individual and team incentives. Indeed, individual incentives schemes in an abstract effort setting automatically lead subjects to choose the efficient level of effort e^*. Instead, the authors study the more interesting case of individual wage-cum-supervision mechanism. In that case, subjects knew that their decision number (abstract effort) was going to be checked with a certain probability p. If their decision number was checked and was below e^* then they received the low wage. Otherwise, they received the high wage.

[12] Circumstantially, the proportion of their time subjects dedicated to Internet under team incentives (28.5%) was remarkably 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)).

We reject the hypothesis that Internet usage was identical for individual and team incentives (see Table A.1 in the appendix). In addition, Internet usage was significantly lower under individual incentives for each of the five periods analyzed separately (see Table A.2 in the appendix). These findings are in line with Hypothesis 1. Also, the positive trend for Internet usage was significantly more pronounced in the team incentives treatment compared with individual incentives (p-value = 0.0142). This suggests that the treatment effect became stronger over time as subjects' fatigue and boredom set in (see Table III.4 in the online appendix).

We summarize our findings regarding Internet usage as follows:

RESULT 2 (Internet usage: Individual versus team incentives).

i) Internet usage was significantly lower in the individual incentives treatment compared with team incentives. This result also holds when analyzing each period separately.

ii) The increase in Internet usage over time was significantly more pronounced in the team incentives treatment than in the individual incentives treatment.

Ours is the first experiment to report a precise measurement of on-the-job leisure activities and demonstrate their significance. Related experiments have stressed the relevance of off-the-job leisure activities that were assessed by analyzing quitting behaviors (Dickinson (1999), Falk and Huffman (2007)) but no studies have attempted to evaluate the importance of on-the-job leisure in a controlled environment.

This finding emphasizes that, in an environment with a long and real-effort task in which fatigue was likely to set in, high-powered incentives were very effective in bringing down Internet usage. Indeed, subjects spent almost three times as long on Internet under team incentives than under individual incentives. This result is consistent with incentives theory (see Holmström (1979), and Laffont and Martimort (2002) for a review).

The introduction of Internet as an alternative activity is a crucial feature of our environment that may have led subjects to consider leisure activities to be as salient as the *work task*. Yet, many subjects never consulted the Internet (40.9% and 11.7% under individual and team incentives, respectively) focusing exclusively on completing the *work task*. Interestingly, we show that the incentives effects identified in the current study largely vanish if we consider an experimental environment in which subjects do not have access to on-the-job leisure activities. In that case, the *work task* is the only activity available to subjects (see part IV-A of the online appendix).

Finally, we expect browsing the Internet and working on the task to compete for the attention of the subjects implying a negative relationship between individual production and Internet usage. We confirm this conjecture by reporting highly significant (p-value < 0.001) negative correlation coefficients between individual production and Internet usage for treatment *I* (-0.67) and treatment *T* (-0.56), respectively. Note that in addition to the *work task* and *Internet browsing*, subjects could obtain earnings from the *clicking task*. As we should expect, no significant differences were observed across treatments in this low-effort task. Subjects successfully clicked on the box in 98% (97%) of its appearances under individual (team) incentives (see Table III.5 in the online appendix).

We introduced peer monitoring in our experimental setting in order to assess whether the shirking behaviors observed under team incentives could be reduced.

3.2. *Peer Monitoring*

We start the analysis of peer monitoring by providing general statistics on watching activities for both monitoring treatments, with (treatment *TP*) and without notification (treatment *TPN*).

3.2.1. *Watching Activities*

Subjects were watched 22.4% (29.9%) of the time in treatment *TP* (*TPN*) while subjects' dedication to monitoring activities was limited to 4.4% (5.3%) of their available time. This occurred because most watchers, regardless of the monitoring treatment, decided to monitor all subjects at the same time. As a result, the amount of time subjects were being watched during the experiment was similar across subjects. In particular, subjects with different levels of performance were being watched for the same amount of time. It was not the case that either low- or high- performers were more likely to be watched by others, regardless of the monitoring treatment. We support this claim by means of a regression analysis in which we introduce as dependent variables the amount of time subjects were watching others and the amount of time they were being watched by others (see Table III.6 in the online appendix).

On average, subjects monitored their peers 5.7 (6.9) times during the experiment for an average duration of 46 (45) seconds per watching episode in treatment *TP* (*TPN*). It is interesting to note that subjects were willing to dedicate a significant amount of their time to monitoring others even in the case in which monitors could not exert peer pressure on other subjects (treatment *TPN*). This suggests that, besides exerting peer pressure, subjects monitor others to obtain feedback about their relative performance as well as to scrutinize others' behavior in the organization.

Comparing monitoring treatments, we observe no significant differences regarding the amount of time subjects spent watching others (see Table A.3 in the appendix). However, we find that subjects were watched significantly more often under the invisible monitoring treatment than under the peer pressure treatment (see Table A.3). This follows from the fact that in the invisible monitoring treatment subjects were significantly less likely to watch only a subset of the other nine organizational members (5.1% of the watching episodes) than under peer pressure (11.1% of the watching episodes).[13] These findings are consistent with the fact that, in treatment *TPN*, monitoring is driven by the willingness to observe others' behaviors and compare oneself with the group while in treatment *TP* monitoring could be partly driven by concerns for exerting peer pressure on all or a subset of the organizational members.

Interestingly, monitoring did not fade away over time. The proportion of their time subjects spent watching others was equal to 4.2% (4.8%) in the first period compared to 5.6% (5.4%) in the last period in treatment *TP* (*TPN*). We cannot reject the hypothesis that the amount of time subjects dedicated to watching activities was the same across periods in treatment *TP* and in treatment *TPN*.[14]

Also, the proportion of subjects who did not watch any other subject did not increase over time and remained constant at a value close to one-third in both monitoring treatments. Considering the experiment as a whole, only 7 out of 60 (3 out of 60) of the subjects did not spend any time monitoring their peers in treatment *TP* (*TPN*) (p-value = 0.186 for a comparison of proportions across treatments). In our experiment, monitoring entailed an opportunity cost since subjects who watched others had to leave the *work task* screen affecting their production negatively. However, these monitoring costs were shared among team members because subjects were paid according to team incentives. As a result, any decline in production due to monitoring activities would affect all workers in the same magnitude. Our environment differs from the model presented by Alchian and Demsetz (1972) and tested by Grosse, Putterman and

[13] In our setting, monitoring all subjects could be done at no extra cost by clicking on the *monitor all* button.

[14] We ran Wilcoxon signed-rank tests to compare the average amount of time subjects spent watching others in each period. To avoid clustering issues we analyzed watching activities at the session level. We ran a total of ten tests for each treatment *TP* and *TPN*, and no p-values were below 0.15, except for the tests comparing average watching times between Periods 1 and 2 (p-value = 0.094) and Periods 2 and 5 (p-value = 0.094) for treatment *TP* and between Periods 2 and 3 (p-value = 0.031) for treatment *TPN*, giving weak support for the fact that subjects watched others more on average in later periods.

Rockenbach (2011) in which subjects who are paid according to their individual contribution would incur an individual cost for undertaking monitoring activities.[15]

Interestingly, subjects rarely watched the same person at the same time. This occurred only in 16.7% and 17.4% of the watching episodes in treatments *TP* and *TPN*, respectively (proportion test, p-value=0.924). It is then not surprising to report that all subjects were watched during the experiment for at least 12 minutes (16 minutes) in treatment *TP* (*TPN*). Finally, we analyze whether the pattern of watching activities within a period followed a random pattern by using a random order test.[16] We analyze for each period of each of the six sessions in each monitoring treatment (that is a total of 30 observations per treatment) whether the order of watching times followed a random order. We find that 25 (28) out of 30 periods were characterized by random watching times in treatment *TP* (*TPN*). We summarize our findings regarding watching activities as follows:

RESULT 3 (Watching activities)

i) Regardless of the monitoring treatment, watching activities were limited to a small percentage of subjects' available time. Nevertheless, all subjects were being watched during the experiment for an average of 22.4% and 29.9% of the duration of the experiment in treatments TP and TPN, respectively.

ii) Regardless of the monitoring treatment, watching activities did not fade away across periods.

iii) Regardless of the monitoring treatment, the pattern of watching activities followed a random order.

iv) The magnitude of watching activities was similar across monitoring treatments. However, monitors were more likely to watch only a subset of subjects in treatment TP than in treatment TPN. As a result, subjects were more likely to be watched in treatment TPN than in treatment TP.

3.2.2. *Comparison of Individual Production Across Treatments*

Similarly to previous treatments, we find that individual production in the monitoring treatments increased over time as is illustrated in Figure 2. We confirm the increase in

[15] Public good games with punishments also consider the case in which the cost for sanctioning other subjects is fully incurred by the individual punisher (Fehr and Gächter (2000)).

[16] We use the random order test in STATA and we consider that the pattern of watching in a given period is not random whenever the test rejects the null hypothesis at a 5% significance level (Swed and Eisenhart (1943)). To that end, we define an indicator variable that takes value one if a subject was watching others in a given minute of a given period and takes value zero otherwise.

production across periods by running a regression of individual production on period dummies and a trend (see Table III.7 in the online appendix).

More importantly, we find that the peer pressure treatment is characterized by significantly higher levels of production than the team incentives treatment while no significant differences are found between peer pressure and individual incentives treatments (see Table A.1 in the appendix). Average production was 47.1% larger in the peer pressure treatment than in the team incentives treatment. Additionally, average total production under peer pressure (20.6) was remarkably close to the case of individual incentives (21.0). These results are in line with Hypothesis 2i. At the same time, we find no support for Hypothesis 2ii according to which production levels in treatment *TP* may surpass those achieved in the individual incentives treatment. Notice that our results hold not only for total production but also for each period analyzed separately (see Table A.2 in the appendix). We conclude that the effect of peer pressure did not vanish across periods since average production (4.62) in the last period was significantly greater in the peer pressure treatment than in the team incentives treatment (3.28). Additionally, average production was 47.3% higher in the last two periods and 46.7% higher in the first two periods in the peer pressure treatment than in the team incentives treatment.

In contrast, invisible monitoring did not lead to any significant increase in either total or period production with respect to the team incentives treatment without monitoring (see Figure 2). At the same time, invisible monitoring led to average production levels which were significantly lower than in the peer pressure and individual incentives treatments. These findings are consistent with Hypotheses 2 and 3. We summarize our findings in Result 4 below.

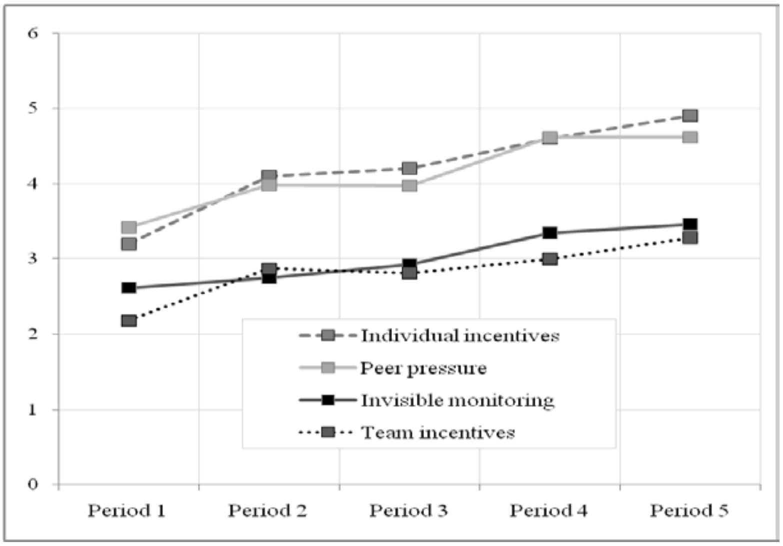

Figure 2. Average production per period for all treatments.

RESULT 4 (*Work task* production: Peer monitoring versus team and individual incentives)

i) Total production was significantly greater in the peer pressure treatment than in the team incentives treatment. This result also holds when analyzing each period separately so that positive peer effects did not vanish over time.

ii) Total production was not significantly different between the peer pressure and the individual incentives treatments. This result also holds when analyzing each period separately.

iii) Total production was not significantly different between the invisible monitoring treatment and the team incentives treatment without monitoring. This result also holds when analyzing each period separately.

iv) Total production was significantly lower in the invisible monitoring treatment than in the peer pressure treatment. This result mostly holds when analyzing each period separately.

The introduction of peer pressure in our experimental design appeared to be a very effective tool that permitted organizations using team incentives to reach efficient levels of production.[17] This result is practically relevant for managers who usually possess limited information about individual contributions, and as a result, cannot rely on individual incentives schemes. The absence of any positive effect on production levels in the invisible monitoring treatment suggests that social pressure is a crucial element of the effectiveness of the monitoring technology. We conducted additional analyses and showed that being watched by others in a given time span of five minutes increased one's own production in the next five to ten minutes in the peer pressure

[17] We interpret the level of production obtained under individual incentives as the efficient level.

treatment (see Table III.8 in the online appendix). At the same time, watching others in a given time span of five minutes did not affect one's own production in the following minutes.

Evidence of positive peer effects has been identified in field studies (Falk and Ichino (2006), Mas and Moretti (2009)), but none of these works have examined peer monitoring as a mechanism to resolve free riding in teams. It is also interesting to observe that we obtained strong peer monitoring effects under anonymity and in the absence of monetary punishments. It is indeed well known that punishments can be very effective in increasing contributions in public good games (Fehr and Gächter (2000), Masclet et al. (2003), Sefton, Shupp and Walker (2007)). Notice that in field studies such as the one designed by Mas and Moretti (2009), workers were not anonymous and could potentially face retaliation for non-cooperative behaviors. In our design, the interaction between subjects was anonymous so as to prevent any form of retaliation after the experiment. In contrast to field studies (Falk and Ichino (2006), Mas and Moretti (2009)) and public good games with threats (Masclet et al. (2003)), subjects were not allowed to communicate in our experiment. The effectiveness of our peer monitoring technology did not rely on physical proximity, verbal threats or face to face interactions. Instead, subjects remained seated at their workstation while monitoring others. Supervisees simply received a notification on their screen that they were currently being watched by another subject (treatment *TP*). The fact that our monitoring technology was highly effective despite the absence of physical proximity and face to face communication is especially relevant given the growing interest for virtual monitoring devices within firms. A large number of programs such as *Spectorsoft*, *Virtual Monitoring*™, *Employee Monitoring* or *Webwatcher* are already available to monitor employees' activities. An early account of computer-based monitoring systems was considered in Chalykoff and Kochan (1989). Our findings are in line with empirical evidence suggesting that mutual monitoring in work groups has been a decisive factor in the success of low-powered firm-wide incentives schemes as is described in the case of Continental Airlines (Knez and Simester (2001)).

In contrast to other supervision mechanisms, peer monitoring does not seem to induce crowding-out of effort which has been reported in recent experimental works (Dickinson and Villeval (2008), Falk and Kosfeld (2006), Frey (1993)). These authors stress that supervision may be perceived as a signal of distrust and, as a result, undermine workers' effort. Frey (1993) as well as Dickinson and Villeval (2008) put forward that the crowding-out effect that results from monitoring activities dominates its disciplining effect when there exist interpersonal relationships between managers and employees whereas the opposite tends to be true in the absence of such relationships. In that respect, our findings are consistent with the works of Frey (1993) and Dickinson and Villeval (2008) since our experimental design is characterized by the absence of interpersonal relationships among workers. Also, crowding-out effects are likely to be more relevant in an organizational structure characterized by a hierarchy. In the current setting, we consider the case of an organization without hierarchy in which all workers had the same roles.

3.2.3. *Comparison of Internet Usage Across Treatments*

Peer monitoring had a considerable impact on Internet usage (see Figure 3).[18] The average proportion of time subjects spent on Internet was significantly lower in the peer pressure treatment (13.1%) than in the team incentives treatment (28.5%). This difference in Internet usage was significant whether considering total Internet usage or Internet usage per period (see Tables A.1 and A.2 in the appendix). Interestingly, we find slightly significant differences in Internet usage between the invisible monitoring treatment (19.8%) and the team incentives treatment (28.5%). This supports the conjecture that social pressure may not be fully eliminated in the invisible monitoring treatment. Subjects may refrain from using the Internet so as to avoid being caught by an invisible monitor. Nevertheless, Internet usage was significantly lower in the peer pressure treatment than under invisible monitoring.

The evolution of Internet usage was remarkably similar for the peer pressure and the individual incentives treatments (see Tables A.1 and A.2 in the appendix). By contrast, internet usage was significantly higher under invisible monitoring (19.8%) than under individual incentives (11.9%).

Similarly to previous treatments, we identify a positive trend in Internet usage for both monitoring treatments (see Table III.9 in the online appendix). The proportion of time subjects dedicated to Internet in treatment *TP* [*TPN*] in the first two periods was only 7.7% [9.7%] on average compared with 16.7% [26.4%] in the last three periods.

[18] The results reported in this section are similar if we analyze working time (time spent on the *work task*) rather than Internet usage (see Tables A.1 and A.2). Using working time instead of Internet usage allows us to control for the fact that monitoring activities may have been used by subjects as an alternative leisure activity. Indeed, one may argue that the low Internet usage in peer monitoring treatments is due to the substitutability between monitoring and Internet activities.

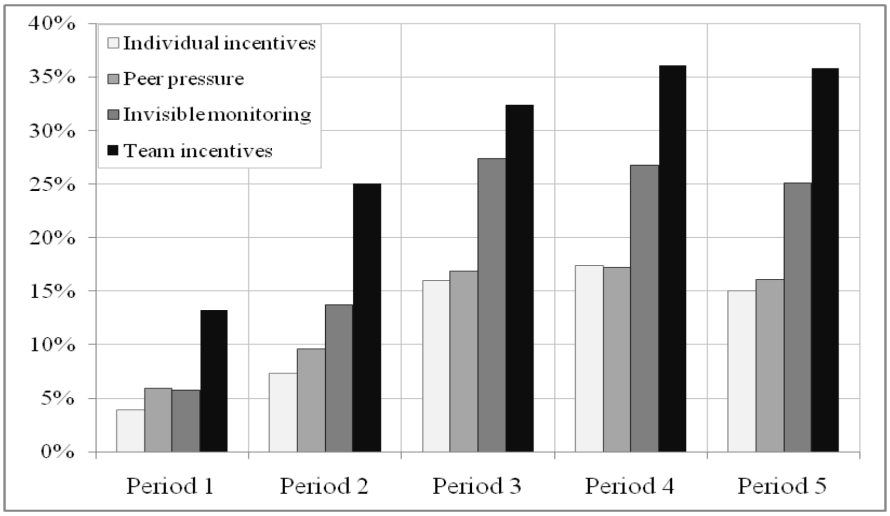

Figure 3. Average Internet usage (in %) for all treatments across periods.

In addition, we analyze whether knowing that they were being watched affected subjects' use of the Internet during the experiment. In particular, we show that in the peer pressure treatment, subjects were less likely to switch from the work task to the Internet if they had been watched by others in the previous five minutes (see Table III.10 in the online appendix). In sum, the introduction of peer monitoring in our experimental design brings down Internet usage. This is an important finding given the growing concern for cyber-slacking (Malachowski (2005), Young (2006)). We summarize our findings as follows:

RESULT 5 (Internet usage: Peer monitoring versus team and individual incentives)

i) Internet usage was significantly lower in the peer pressure treatment than in the team incentives treatment. Also, Internet usage was marginally higher in the invisible monitoring treatment than in the team incentives treatment.

ii) Internet usage was not significantly different between the peer pressure and the individual incentives treatments. However, Internet usage was significantly higher in the invisible monitoring treatment than in the individual incentives and peer pressure treatments.

4. CONCLUSIONS

In this paper, we develop a software for the analysis of organizational issues in the laboratory. The design of such experimental environment may be seen as a pertinent element in the development of Experimental Organizational Economics (Camerer and Weber, 2012, p.215). In particular, we incorporated several features of existing firms in a virtual organization by allowing subjects to allocate their time between a real-effort task that created value for the organization and a real-leisure activity. We considered the most decentralized form of organizations in which no hierarchies existed and all subjects had the same role. This represented a natural starting point in our effort to identify elements that lead to organizational success.

As a first step, we compared organizations using team and individual incentives in order to assess the relevance of incentives effects in our virtual organizations. We found that individual incentives led to significantly higher levels of production and significantly lower levels of Internet usage than team incentives. These findings confirmed that implementing high-powered incentives schemes is an important factor of organizational success consistently with theoretical research (Holmström (1979), see Laffont and Martimort (2002) for a review).

We studied peer monitoring as an example of mechanism that may allow organizations to recover the efficiency loss provoked by the use of weak incentives. We found that using peer monitoring devoid of punishment in combination with team incentives allowed organizations to reach production levels that were as high as in the case of individual incentives. In contrast to public good games with punishments, both effort and efficiency were increased by the use of peer monitoring. To our knowledge, ours is the first controlled experiment showing that peer monitoring can offset the loss in efficiency resulting from the use of low-powered incentives schemes. Peer monitoring was particularly effective because subjects spent a limited amount of time watching others while sharing the monitoring burden so that all subjects were being watched at least once during the experiment. It is as if people possessed natural skills for peer monitoring and understood both its positive effect on productivity as well as the negative consequences of its intensive use.

This is good news for most organizations that cannot rely on precise measures of individual contributions. Peer monitoring is traditionally seen as a decisive advantage of organizations where its effectiveness usually relies on face to face and repeated interactions among parties that are inherent to the organizational environment (Bandiera, Barankay and Rasul (2005), Falk and Ichino (2006), Mas and Moretti (2009)). Interestingly, the implementation of virtual monitoring devices of the type used in our study may mitigate the comparative advantage of traditional organizations vis-à-vis virtual organizations and other decentralized organizational structures.

REFERENCES

Abbink, K., Brandts, J., Herrmann, B, and H. Orzen (2010). "Intergroup Conflict and Intra-group Punishment in an Experimental Contest Game," *American Economic Review*, 100(1), 420-47.

Alchian, A., and H. Demsetz (1972), “Production, information costs, and economic organization,” *American Economic Review*, 62, 777-795.

Andreoni, J., And B. D. Bernheim (2009): “Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects,” *Econometrica*, 77(5), 1607-1636.

Azmat, G., and N. Iriberri (2010): “The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students,” Journal of Public Economics, 94(7-8), 435-452.

Babcock, P., Bedard, K., Charness, G., Hartman, J., and H. Royer. (2012). “Letting Down the Team? Evidence of Social Effects of Team Incentives,” Working Paper UCSB.

Bandiera, O., I. Barankay, and I. Rasul (2005): “Social Preferences and the Response to Incentives: Evidence from Personnel Data,” *Quarterly Journal of Economics*, 120(3), 917-962.

Bandiera, O., I. Barankay, and I. Rasul (2012): “Team Incentives: Evidence from a Field Experiment,” forthcoming *Journal of the European Economic Association*.

Barnard, C. I. (1938): The functions of the executive. Cambridge, MA: Harvard University Press.

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

Barzel, Y. (1982): “Measurement Costs and the Organization of Markets,” *Journal of Law and Economics*, 25(1), 27-48.

Blanes i Vidal, J., and M. Nossol (2011): “Tournaments without Prizes: Evidence from Personnel Records,” *Management Science*, 57, 1721-1736.

Bucciol, A., Houser, D., and M. Piovesan (2011): “Temptation at Work,” Working Paper 11-090, Harvard Business School.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2010): “Overconfidence is a Social Signaling Bias,” mimeo.

Camerer, C., and R. Weber. (2012). “Experimental Organizational Economics,” R. Gibbons and J. Roberts (eds.), Handbook of Organizational Economics. Princeton, NJ: Princeton University Press, Chapter 6, 213-262.

Carpenter, J. P. (2007a): “Punishing Free Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods,” *Games and Economic Behavior*, 60(1), 31-51.

Carpenter, J. P. (2007b): “The demand for punishment,” *Journal of Economic Behavior and Organization*, 62, 522–542.

Carpenter, J., S. Bowles, and H. Gintis (2006): “Mutual Monitoring in Teams: Theory and Experimental Evidence on the Importance of Reciprocity,” *IZA DP*, 2106.

Carpenter, J. P., S. Bowles, H. Gintis, and S. H. Hwang (2009): “Strong Reciprocity and Team Production: Theory and Evidence,” *Journal of Economic Behavior and Organization*, 71(2), 221-232.

Chalykoff, J., and T. Kochan (1989): “Computer-aided Monitoring: Its Influence on Employee Job Satisfaction and Turnover,” *Personnel Psychology*, 42, 807-834.

Chandler, A.D. (1992): “Organizational Capabilities and the Economic History of the Industrial Enterprise,” *Journal of Economic Perspectives*, 6(3), 79–100.

Charness, N., and J. I. D. Campbell (1988): “Acquiring Skill at Mental Calculation in Adulthood: A Task Decomposition,” *Journal of Experimental Psychology: General*, 117, 115-129.

Charness, G., and P. Kuhn (2011): “Lab Labor: What Can Labor Economists Learn from the Lab?,” in Handbook of Labor Economics, 4, ed. by O. Ashenfelter and D. Card. Elsevier, chapter 03, 229-330.

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.

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

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 (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.

Dumaine, B. (1990): "Who needs a boss?," *Fortune*, 121, 52-60.

Dumaine, B. (1994): "The Trouble with Teams," *Fortune*, 130(5), 86-92.

Encinosa, W., M. Gaynor, and J. Rebitzer (2007): "The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems," *Journal of Economic Behavior and Organization*, 62, 187-214.

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 U. Sunde (2006): "Self-Confidence and Search," IZA Discussion Paper.

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

Falk, A., and M. Kosfeld (2006): "The Hidden Costs of Control," *American Economic Review*, 96, 1611-1630.

Farrell, G. (2000): "Online Time Soars at Office; Not All Surfing Work-Related," *USA Today*, February 18, http://www.usatoday.com/life/cyber/tech/cth391.htm?loc=interstitialskip.

Fehr, E., and S. Gächter (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90(4), 980-994.

Fehr, E., A. Klein, and K. M. Schmidt (2007a): "Fairness and Contract Design," *Econometrica*, 75(1), 121-154.

Fehr, E., A. Klein and K. M. Schmidt (2007b): "Adding a Stick to the Carrot? The Interaction of Bonuses and Fines," *American Economic Review papers and proceedings*, 97(2), 177-181.

Frey, B. (1993): "Does Monitoring Increase Work Effort? The Rivalry with Trust and Loyalty," *Economic Inquiry*, 31, 663-70.

Gneezy, U. and List, J. A. (2006), "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," *Econometrica*, 74(5), 1365–1384.

Gordon, R. J. (2000): "Does the "New Economy" Measure up to the Great Inventions of the Past?," *The Journal of Economic Perspectives*, 14(4), 49-74.

Grosse, S., L. Putterman, and B. Rockenbach (2011): "Monitoring in Teams: A Model and Experiment on the Central Monitor Hypothesis," *Journal of the European Economic Association*, 9(4), 785-816.

Hamilton, B., J. Nickerson, and H. Owan (2003): "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 111(3), 465-497.

Hansen, D. G. (1997): "Worker Performance and Group Incentives: A Case Study," *Industrial and Labor Relations Review*, 51(1), 37-49.

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

Holmström, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324-340.

Houser, D., Schunk, D., Winter, J., and E. Xiao. (2010), "Temptation and Commitment in the Laboratory," Institute for Empirical Research in Economics, University of Zurich Working Paper 488.

Ichniowski, C., T. A. Kochan, D. Levine, C. Olson, and G. Strauss (1996): "What Works at Work: Overview and Assessment," *Industrial Relations: A Journal of Economy and Society*, 35(3), 299-333.

Ichniowski, C., K. Shaw, and G. Prennushi (1997): “The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines,” *American Economic Review*, 87(3), 291-313.
Jensen, M. C., and W. H. Meckling (1976): “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure,” *Journal of Financial Economics*, 3(4), 305-360.
Jones, S., C. Johnson-Yale, S. Millermaier, and F. S. Perez (2009): “Everyday Life, Online: U.S. College Students' Use of the Internet,” *First Monday*, 14(10),
Kandel, E., and E. Lazear (1992): “Peer Pressure and Partnerships,” *Journal of Political Economy*, 100, 801-817.
Knez, M. and D. Simester (2001): “Firm-wide Incentives and Mutual Monitoring at Continental Airlines,” *Journal of Labor Economics*, 19(4), 743–772
Kruse, D. L. (1992): “Profit Sharing and Productivity: Microeconomic Evidence from the United States,” Economic Journal, 102(410), 24-36.
Kuhnen, C. M., and A. Tymula (2009): “Rank Expectations, Feedback and Social Hierarchies,” mimeo.
Laffont, J. J., and D. Martimort (2002): The Theory Of Incentives: The Principal-Agent Model. Princeton University Press: Princeton, N.J.; Oxford.
Ledyard, J. (1995): Public Goods: A Survey of Experimental Research, in Handbook of experimental economics, ed. by J. H. Kagel and A. E. Roth. Princeton: Princeton University Press, 111-94.
Leibowitz, A., and R. Tollison (1980): “Free-Riding, Shirking and Team Production in Legal Partnerships,” *Economic Inquiry*, 18, 380-394.
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
Manz, C. C., and H. P. Sims Jr (1993): Business without Bosses. New York: Wiley.
Mas, A., and E. Moretti (2009): “Peers at Work,” *American Economic Review*, 99(1), 112–145.
Masclet, D., C. Noussair, S. Tucker, and M. C. Villeval (2003): “Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism,” *American Economic Review*, 93(1), 366-380.
Nalbantian, H., Schotter, A. (1997): “Productivity Under Group Incentives: An Experimental Study,” *American Economic Review*, 87(3), 314-341.
Niederle, M., and L. Vesterlund (2007): “Do Women Shy Away from Competition? Do Men Compete Too Much?,” *The Quarterly Journal of Economics*, 3(8), 1067-1101.
Nikiforakis, N. (2010). “Feedback, Punishment and Cooperation In Public Good Experiments,” *Games and Economic Behavior*, 68, 689-702.
Sacerdote, B. (2001): “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *Quarterly Journal of Economics*, 116(2), 681-704.
Sefton, M., R. Shupp, and J. Walker (2007): “The Effect of Rewards and Sanctions in Provision of Public Goods,” *Economic Inquiry*, 45, 671-690.
Swed, F., and C. Eisenhart (1943): “Tables for Testing Randomness of Grouping in a Sequence of Alternatives,” *Annals of Mathematical Statistics*, 14, 66-87.
Van Dijk, F., J. Sonnemans, and F. van Winden (2001): “Incentive Systems in a Real Effort Experiment,” *European Economic Review*, 45, 187-214.
Young, K.S (2005): “Internet Addiction Prevention and Education: Preventive Education Training Can Reduce Problematic Internet Use in the Workplace and Help Employers Create a Work Environment that Encourages Identification and Treatment of Addictive Online Behaviors,” *The Journal of Employee Assistance*, 15(3), 15-19.
Young, K.S (2006): “The Internet and Workplace Transformation,” *Advances in Management Information Systems*, 7, 193-204.
Zajonc, R. B. (1965): “Social Facilitation,” *Science*, 149(3681), 269–74.
Zimmerman, D. J. (2003): “Peer Effects in Academic Outcomes: Evidence From a Natural Experiment,” *Review of Economics and Statistics*, 85(1), 9–23.

6. APPENDIX

Table A.1. P-values for statistical tests assessing differences in production and Internet usage across treatments

		Clustered t-test	Clustered Wilcoxon rank-sum test	t-test	Wilcoxon rank-sum test	t-test (group averages)	Wilcoxon rank-sum test (group averages)
Treatment *I* vs. Treatment *T*	Production	**0.002**	**0.009**	**0.002**	**0.003**	**0.004**	**0.012**
	Internet usage[19]	**<0.001**	**0.008**	**<0.001**	**<0.001**	**0.012**	**0.022**
Treatment *T* vs. Treatment *TP*	Production	**0.019**	**0.049**	**0.008**	**0.015**	**0.046**	**0.045**
	Internet usage	**0.003**	**0.010**	**<0.001**	**<0.001**	**0.019**	**0.041**
	Working time	**0.035**	**0.094**	**0.015**	**0.042**	**0.070**	**0.093**
Treatment *I* vs. Treatment *TP*	Production	0.866	0.712	0.860	0.751	0.832	0.628
	Internet usage	0.728	0.754	0.725	0.828	0.678	0.534
	Working time	0.114	**0.058**	0.118	**0.018**	0.127	0.138
Treatment *I* vs. Treatment *TPN*	Production	**0.009**	**0.024**	**0.009**	**0.013**	**0.018**	**0.022**
	Internet usage	**0.042**	**0.047**	**0.036**	**0.014**	**0.066**	0.101
	Working time	**0.001**	**0.005**	**0.001**	**<0.001**	**0.007**	**0.005**
Treatment *T* vs. Treatment *TPN*	Production	0.687	0.727	0.680	0.738	0.696	0.688
	Internet usage	0.110	0.114	**0.059**	**0.066**	0.149	0.240
	Working time	0.531	0.898	0.463	0.894	0.550	0.394
Treatment *TP* vs. Treatment *TPN*	Production	**0.054**	**0.099**	**0.027**	**0.044**	**0.088**	0.132
	Internet usage	**0.097**	**0.067**	**0.099**	**0.022**	0.129	0.310
	Working time	**0.070**	**0.060**	**0.072**	**0.023**	**0.092**	0.240

[19] Working time p-values are identical to Internet usage, as there are no other activities available in these treatments.

Table A.2. P-values for clustered t-tests (clustered Wilcoxon rank-sum tests) assessing differences in production and Internet usage per period across treatments

		Period 1	Period 2	Period 3	Period 4	Period 5
Treatment *I* vs. Treatment *T*	Production	**0.025** **(0.029)**	**0.026** **(0.028)**	**0.025** **(0.043)**	**0.006** **(0.025)**	**0.004** **(0.020)**
	Internet usage	**0.003** **(0.021)**	**0.002** **(0.019)**	**0.024** **(0.050)**	**0.006** **(0.013)**	**<0.001** **(0.002)**
Treatment *T* vs. Treatment *TP*	Production	**0.012** **(0.028)**	**0.061** **(0.124)**	**0.060** **(0.094)**	**0.039** **(0.095)**	**0.035** **(0.084)**
	Internet usage	**0.034** **(0.128)**	**0.011** **(0.032)**	**0.044** **(0.083)**	**0.008** **(0.021)**	**0.001** **(0.007)**
	Working time	0.383 (0.280)	**0.049** (0.332)	0.144 (0.560)	**0.045** (0.121)	**0.025** **(0.047)**
Treatment *I* vs. Treatment *TP*	Production	0.755 (0.778)	0.900 (0.639)	0.757 (0.819)	0.959 (0.932)	0.581 (0.537)
	Internet usage	0.283 (0.187)	0.438 (0.785)	0.874 (0.490)	0.963 (0.747)	0.842 (0.265)
	Working time	**0.003** **(0.003)**	**0.063** **(0.058)**	0.341 **(0.058)**	0.456 **(0.097)**	0.192 **(0.011)**
Treatment *I* vs. Treatment *TPN*	Production	0.198 (0.185)	**0.007** **(0.020)**	**0.052** **(0.075)**	**0.037** **(0.065)**	**0.012** **(0.028)**
	Internet usage	0.405 (0.230)	**0.046** (0.104)	**0.058** **(0.069)**	0.121 **(0.081)**	**0.072** **(0.021)**
	Working time	**0.007** **(0.017)**	**<0.001** **(0.003)**	**0.005** **(0.008)**	**0.020** **(0.007)**	**0.009** **(0.003)**
Treatment *T* vs. Treatment *TPN*	Production	0.325 **(0.072)**	0.779 (0.987)	0.816 (0.900)	0.652 (0.723)	0.744 (0.812)
	Internet usage	**0.040** (0.129)	**0.073** (0.188)	0.533 (0.390)	0.233 (0.168)	0.102 **(0.022)**
	Working time	0.479 (0.522)	0.436 (0.631)	0.976 (0.468)	0.616 (0.916)	0.418 (0.813)
Treatment *TP* vs. Treatment *TPN*	Production	**0.096** (0.176)	**0.043** **(0.087)**	0.111 (0.122)	0.108 (0.156)	**0.069** **(0.099)**
	Internet usage	0.941 (0.895)	0.265 (0.226)	0.100 (0.160)	0.127 (0.146)	0.106 (0.148)
	Working time	0.883 (0.839)	**0.076** **(0.085)**	**0.072** (0.152)	0.116 **(0.092)**	0.128 (0.105)

Table A.3. P-values associated with the treatment dummy capturing differences across monitoring treatments.

	Watching time	Length of watching episodes	Proportion of watching episodes for which only one (all) subject(s) is monitored	Amount of time being watched	Number of times a subject is watched
Regression type	Tobit with random effects	Tobit with random effects	Probit with random effects	Tobit with random effects	Poisson with random effects
P-value associated with the treatment dummy	0.165	0.573	0.041 (0.075)	0.000	0.000

All regressions are completed at the minute level and all include a trend. These results are robust to the cases of the 5-minute analysis as well as to the case of the analysis per period.

ONLINE APPENDIX

PART I. SCREENSHOTS

	Column1	Column2	Column3	Column4	Column5	Column6	Sum Row:
	3.00	6.00	3.00	0.00	6.00	0.00	
	10.00	5.00	1.00	5.00	2.00	3.00	
	8.00	3.00	5.00	4.00	8.00	7.00	
	1.00	6.00	0.00	9.00	8.00	0.00	
	3.00	7.00	0.00	8.00	10.00	4.00	
	3.00	10.00	10.00	6.00	10.00	0.00	
Sum Column:							

Figure O.1. Example of table summation for the *work task*.

Figure O.2. Embedded Internet screen.

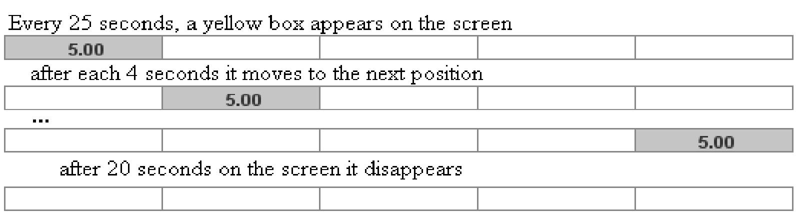

Figure O.3. The *clicking task*.

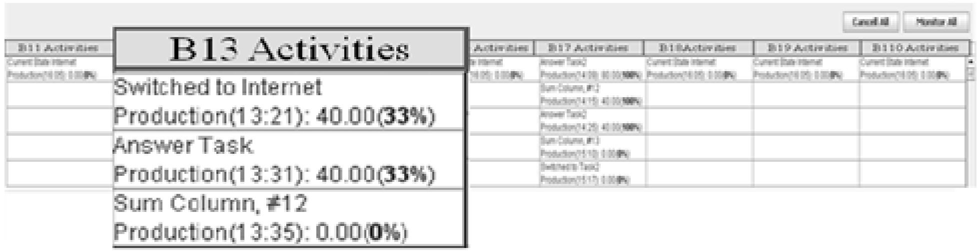

Figure O.4. Monitoring screen with a zoom on subject B13.

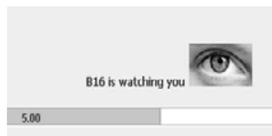

Figure O.5. Notification when a subject is being watched.

PART II. CONCEPTUAL FRAMEWORK

We solve the model by maximizing the utility function of worker $i \in \{1,...,N\}$ with respect to work effort e_i for treatment *I* and treatment *T*. The derivations for equilibrium work effort in the case of treatments *I* and *T* are trivial. For treatments where monitoring is available (*TP* and *TPN*), we also maximize the utility function with respect to m_i as follows:

$$\underset{e_i, m_{i,\mu}}{Max} U_i := s_{i,\mu} \sum_{k=1}^{N} e_{k,\mu} - \alpha \frac{e^2{}_{i,\mu}}{2} - \beta \frac{(e_{i,\mu} + m_{i,\mu})^2}{2} - v_\mu \chi (e_{i,\mu} - e_0)(N-1) m_{j,\mu}$$

s.t $0 \le e_i + m_{i,\mu} \le 1$, where $\mu \in \{TP, TPN\}$ [1]

For $i \neq j$, the first order conditions of [1] are:

$$\frac{\partial U_i}{\partial e_i} = 0 \Leftrightarrow e_\mu^* = \frac{\frac{1}{N} - \beta m_{i,\mu} + v_\mu \chi (N-1) m_{j,\mu}}{\alpha + \beta}$$

$$\frac{\partial U_i}{\partial m_{i,\mu}} = 0 \Leftrightarrow m_{i,\mu}^* = \frac{(N-1)^2}{N} \frac{v_\mu \chi}{\alpha\beta} - \frac{1}{\alpha N} - \frac{v_\mu \chi}{\alpha}(N-1) m_{j,\mu}$$

As a result, in a symmetric equilibrium, it follows that $m_{i,\mu}^* = m_{j,\mu}^* = m_\mu^*$ and $m_\mu^* = \frac{\frac{1}{\alpha N}\left[\frac{v_\mu \chi}{\beta}(N-1)^2 - 1\right]}{1 + \frac{v_\mu \chi}{\alpha}(N-1)}$ and $e_\mu^* = \frac{\frac{1}{N} + [v_\mu \chi (N-1) - \beta] m_\mu^*}{\alpha + \beta}$.

As a result, as long as $v_\mu \chi > \frac{\beta}{N-1}$ then both $m_\mu{}^* > 0$ and $e_\mu{}^* > 0$. Also, there exists a vector of parameters $(\alpha^*, \beta^*, v^*{}_\mu, \chi^*)$ such that an equilibrium exists in which $(m_\mu{}^*, e_\mu{}^*) \in (0,1)$. This is the case as long as $v_\mu^* \chi^* > \frac{\beta^*}{N-1}$ and α^* is sufficiently high. In that case, the work effort under treatment *TP* can be larger than under individual incentives as long as $v_\mu^* \chi^* > \frac{1 - \frac{1}{N} + \beta^* m_\mu^*}{N-1}$.

It follows from equilibrium values of work effort and monitoring that $\frac{\partial m_\mu^*}{\partial \chi} > 0$ and $\frac{\partial e_\mu^*}{\partial \chi} > 0$.

Also, $\frac{\partial m_\mu^*}{\partial \beta} < 0$ and $\frac{\partial e_\mu^*}{\partial \beta} < 0$.

PART III. ADDITIONAL ANALYSES[20]

Table III.1. Tobit regression with random effects for individual production per period (Treatments *I* and *T*)

	Regression 1		Regression 2	
	Treatment I	Treatment T	Treatment I	Treatment T
Intercept	2.542***	1.101**	2.395***	1.044**
Trend	-	-	0.477***	0.336***
Period 2	1.207***	1.003**	-	-
Period 3	1.278***	0.924**	-	-
Period 4	1.711***	1.211***	-	-
Period 5	2.151***	1.592***	-	-
Number of observations and Log likelihood	*n* = 330 66 Left-censored -709.6 (Prob>χ²)<0.001	*n* = 300 101 Left-censored -557.4 (Prob>χ²)<0.001	*n* = 330 66 Left-censored -711.1 (Prob>χ²)<0.001	*n* = 300 101 Left-censored -711.0 (Prob>χ²)<0.001

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

We confirm that period production stagnated in Period 3 as is revealed by comparing the coefficient associated with Period 2 and Period 3 dummies (p-value = 0.8676).

Table III.2. Tobit regression with random effects for individual production per period

	Treatment *T*	Treatment *I*
Intercept	-0.983	2.49***
Trend	0.147	0.336*
Group production in *(t-1)*[21]	0.597*	-0.009
Group production in *(t-1)* × Dummy greater than group average in *(t-1)*[22]	-0.908**	-0.03
Dummy greater than group average in *(t-1)*	4.675***	0.789
Number of observations and Log likelihood	*n* = 240 76 Left-censored -451.4 (Prob>χ²)<0.001	*n* = 264 45 Left-censored -569.6 (Prob>χ²)=0.103

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

[20] Note that the results presented in this section are robust to using linear regressions instead of tobit regressions or using clustered standard errors instead of (or in addition to) random effects. These results are available upon request.
[21] Group production excludes a given subject's individual production so as to avoid endogeneity issues.
[22] This dummy variable takes a value of one if a given subject produces strictly more than the average of the other group members in a given period.

Table III.3. Tobit regression with random effects for Internet usage per period (Treatments *I* and *T*)

	Regression 1		Regression 2	
	Treatment I	Treatment T	Treatment I	Treatment T
Intercept	0.029	0.730	-0.366	0.368
Trend	-	-	0.661***	1.196***
Period 2	0.670	2.463***	-	-
Period 3	2.448***	4.125***	-	-
Period 4	2.766***	4.799***	-	-
Period 5	2.253***	4.786***	-	-
Number of observations and Log likelihood	n = 330 8 Right-censored -921.3 (Prob>χ^2)=0.001	n = 300 23 Right-censored -930.2 (Prob>χ^2)<0.001	n = 330 8 Right-censored -926.5 (Prob>χ^2)<0.001	n = 300 23 Right-censored -933.5 (Prob>χ^2)<0.001

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

Table III.4. Tobit regression with random effects for Internet usage per period in Treatments *I* and *T*

Intercept	0.578
Trend	1.178***
Treatment (Dummy that takes value one if treatment is *I*)	-1.223
Trend×Treatment	-0.510**
Number of observations and Log likelihood	n = 630 31 Right-censored -1882.7 (Prob>χ^2)<0.001

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

Table III.5. Clicking task performance and timing across treatments

Clicking task	Treatment *I*	Treatment *T*	Treatment *TP*	Treatment *TPN*
Success rate	98%	97%	99%	99%
P-value clustered Wilcoxon rank-sum test				
vs. Treatment T	0.616	-	-	-
vs. Treatment TP	0.475	0.784	-	-
vs. Treatment TPN	0.460	0.342	0.579	-

Success rate: Average proportion of the 240 yellow boxes subjects had clicked before they disappear from the screen.

Table III.6. Tobit regressions with random effects for watching activities in a given period as a function of production in the previous period (Treatment *TP*)

Dependent variable:	Time being watched		Time spent watching	
	Treatment *TP*	Treatment *TPN*	Treatment *TP*	Treatment *TPN*
Intercept	21.965***	27.179***	5.903***	6.688***
Individual production in the previous period	0.021	0.181	-0.075	-0.091
No. of observations	*n* = 240 70 Left-censored	*n* = 240 70 Left-censored	*n* = 240 70 Left-censored	*n* = 240 70 Left-censored
Log likelihood	-381.5 (Prob>χ^2)=0.258	-590.9 (Prob>χ^2)=0.570	-434.9 (Prob>χ^2)=0.177	-608.9 (Prob>χ^2)=0.194

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

Table III.7. Tobit regression with random effects for individual production per period (Treatment *TP* and *TPN*)

	Regression 1		Regression 2	
	Treatment *TP*	Treatment *TPN*	Treatment *TP*	Treatment *TPN*
Intercept	2.826***	1.516**	2.609***	1.104*
Trend	-	-	0.356***	0.288**
Period 2	0.675	0.144	-	
Period 3	0.720*	0.182	-	
Period 4	1.433***	0.819**	-	
Period 5	1.405***	1.098***	-	
No. of observations	n = 300 70 Left-censored	n = 300 95 Left-censored	n = 300 70 Left-censored	n = 300 95 Left-censored
Log likelihood	-618.0 (Prob>χ^2)=0.004	-545.3 (Prob>χ^2)=0.026	-618.8 (Prob>χ^2)=0.004	-545.9 (Prob>χ^2)=0.002

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

In Table III.8, we use a 5-minute time frame to assess the impact of watching activities on real-time production. The independent variables related to watching activities are referred to as *Watching* and *Being watched*. These variables measure the amount of time (in seconds) that a subject spent watching others (*Watching*) and the amount of time (in seconds) a subject was watched (*Being watched*) by at least one subject in a given time span of five minutes. We introduce independent variables with lags so as to mitigate possible endogeneity issues.[23] We include a trend as independent variable so as to control for the steady increase of production across periods.[24-25]

[23] Endogeneity issues may arise if we introduce the current amount of time subjects spent watching others as well as the current amount of time they were being watched by others as independent variables. Indeed, one may expect that individual production could cause changes in watching behaviors. For example, subjects with low levels of production may feel ashamed (Kandel and Lazear (1992)) and decide to avoid consulting the performance of others.

[24] Similar results are obtained when controlling for beginning or end of period effects. For example, the nature of our results is unchanged when introducing in our regression analysis a dummy variable that takes value one if the five minute time span corresponds to the first (last) five minutes of the period.

[25] A number of other specifications have been considered such as including up to three lags in the independent variables or adding group production in the previous period as regressors. These specifications gave similar results. We also used dynamic panel data models with Arellano-Bond (1991) estimation technique. However, this estimation technique was not successful in fully eliminating residual autocorrelation as we may expect given the limited number of instrumental variables at our disposal. Finally, note that the results are robust to using linear regressions instead of tobit regressions or using clustered standard errors instead of (or in addition to) random effects. These results are available upon request.

Table III.8.[26] Tobit regression with random effects for individual production in a 5-minute time span

	Coefficients
Intercept	-0.803***
Being watched in *t-1*	0.001
Being watched in *t-2*	0.002***
Watching in *t-1*	-0.001
Watching in *t-2*	0.001
Trend	0.066***
Number of observations and Log likelihood	*n* = 1080 398 left-censored Log likelihood = -1324.595, Prob > χ^2 = 0.000

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

The delay in the impact of watching activities can be accounted for by the time subjects needed to produce a table (4 minutes and 3 seconds on average) in order to increase their individual production. Furthermore, subjects who responded positively to peer monitoring by switching from *Internet browsing* to the *work task* may have needed an additional amount of time to return their concentration to the *work task*.

Table III.9. Tobit regression with random effects for Internet usage per period (Treatment *TP* and *TPN*)

	Regression 1		Regression 2	
	Treatment *TP*	Treatment *TPN*	Treatment *TP*	Treatment *TPN*
Intercept	1.191*	1.178	0.657	0.859
Trend	-	-	0.563***	1.056***
Period 2	0.722	1.600**	-	
Period 3	2.188***	4.425***	-	
Period 4	2.245***	4.279***	-	
Period 5	2.019***	3.930***	-	
No. of observations Log likelihood	*n* = 300 2 Right-censored -813.4 (Prob>χ^2)<0.001	*n* = 300 12 Right-censored -900.7 (Prob>χ^2)<0.001	*n* = 300 2 Right-censored -818.7 (Prob>χ^2)<0.001	*n* = 300 12 Right-censored -907.3 (Prob>χ^2)<0.001

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

We also confirm that period production stagnated in Period 3 as is revealed by comparing the coefficient associated with Period 2 and Period 3 dummies (t-test, p-value=0.916 and p-value=0.926 for Treatments *TP* and *TPN*, respectively).

[26] An independent variable accounting for the number of watchers is not statistically significant when introduced in the specification of the regression. This may be due to the fact that the information on the number of watchers was not made particularly salient. In case a subject was watched by more than one person, the following indication was printed on the screen: "more than one subject is watching you".

We assessed whether being watched in a given 5-minute time span led subjects to switch to the *work task* screen in the following five minutes. We run a Logistic regression where the dependent variable *Work task* is a dummy variable that takes value one if the corresponding subject was on the *Work task* screen in a given 5-minute time span and zero otherwise (see Table III.10). We find that, the more time a subject was being watched in a given 5-minute time span the more likely he or she was to be on the *work task* screen in the following five minutes. This is the case since the coefficient associated with *Being watched* is positive and significant.

Table III.10. Logistic regression with random effects (Treatment *TP*)

Dependent variable:	*Work task* in *t*
Intercept	5.702 ***
Being watched in *t-1*	0.008***
Individual production in t-1	0.769***
Trend (Period)	-0.601***
Dummy Minute 6 to 10[27]	0.148
Dummy Minute 11 to 15	-0.893*
Dummy Minute 16 to 20	-1.191**
No. of observations	n = 1140
Log likelihood	-172.568 (Prob>χ^2)<0.001

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

Reference

Arellano, M., and S. Bond (1991): “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *The Review of Economic Studies*, 58, 277-297.

[27] The *Minute* variables are dummy variables that take value one for a given time frame of five minutes. We use dummy variables for each time frame of five minutes so as to control for the rising use of Internet within a given period.

PART IV. ROBUSTNESS CHECKS

A- INCENTIVES EFFECTS IN THE ABSENCE OF INTERNET BROWSING

We conducted two additional treatments so as to assess the robustness of the incentives effects established in Result 1 according to which team incentives underperform individual incentives. To that end, we recruited a total of 127 subjects to compare production levels under team incentives (60 subjects) and individual incentives (67 subjects) in a context in which Internet browsing was not available to subjects. We find that incentives effects were much more limited in the case in which Internet browsing was not available compared with the case in which it was available. In particular, in the absence of the Internet browsing option, average production levels were only 28.1% larger under individual incentives (21.8) than under team incentives (17.0). These differences are only marginally significant.[28] Remember that in the case in which Internet browsing was available, differences in average production levels were substantial (48.8%) and highly significant (p-value < 0.01). These results suggest that introducing Internet browsing as an on-the-job leisure alternative is a crucial feature of our environment that allows us to identify incentives effects which may otherwise be largely underestimated or even fail to be observed (Dohmen and Falk (2011), van Dijk, Sonnemans and van Winden (2001)).

B- SURVEY

We conducted robustness checks for our treatment effects controlling for subjects' ability levels and demographic information such as age, gender and working experience. To do so, we invited our subjects to come back to the laboratory and participate in a follow-up study in which we measured subjects' arithmetic skills and gathered demographic information. We invited the 186 subjects who were involved in treatments *I*, *T* and *TP* to participate in a one-hour survey.[29] A total of 111 participants (60% of the initial sample) ranging from 18 to 28 years old (mean= 20.35, s.d =1.97) came back to answer the survey, 52 of which were female. The subjects who came back were distributed across the three treatments (37, 31 and 43 students in treatments *I*, *T* and *TP*, respectively).

[28] The p-values for the Wilcoxon (clustered) rank sum test were equal to 0.0734 (0.0967) while p-values for the t-test (clustered t-test) were equal to 0.0322 (0.0307).
[29] These invitations were sent on average three months after subjects' initial participation in any of the three treatments.

No significant differences were found in average production levels between the subset of subjects who came back for the follow-up survey (19.6) and the subjects (17.6) who did not come back. This is shown in the following regression.

Table IV.1. Tobit regression with random effects.
(Treatment *TP*)

Dependent variable:	*Total production*
Intercept	16.394***
Survey Dummy	1.972
Number of observations	n = 186
Log likelihood	-681.079 (Prob>χ^2)=0.389

Survey Dummy takes value 1 if a subject participated in the follow-up survey, and 0 otherwise.
*p -value<.10, ** p-value<.05, and *** p-value<.01.

In the subset of subjects who came back for the follow-up study, the proportion of subjects who were in the top three performers of their experimental session (30 out of 111) in the initial experiment was the same as the proportion of subjects who were in the bottom three (30 out of 111). Also, these proportions were not different from 30% which is the proportion of top three and bottom three subjects in our initial experiments (χ^2 = 1.635, p-value = 0.441).

In the survey, subjects had to answer questions related to demographics, personality traits and arithmetic skills. In particular, following Dohmen and Falk (2011), we measured subjects' summation skills in a five-minute incentivized exercise which was similar to the *work task* in the current experimental design.[30]

We categorized subjects as low-ability workers whenever their performance on the arithmetic task was lower than the median performance (n=56) while high-ability workers were characterized by performance levels which were greater than or equal to the median performance (n=55). We then define *Arithmetic skills* as a dummy variable that takes value 1 if a subject is categorized as being a high-ability worker and takes value 0 otherwise. We include treatment dummies in our regression which take value 1 if a subject had previously participated in the corresponding treatment and value 0 otherwise. We were able to confirm the significance of our treatment effects (see Table IV.2).

[30] Subjects earned lottery tickets according to their performance on the task. A lottery-prize of 400$ was paid to a single winner among all participants. The lottery prize was known to all participants.

TABLE IV.2. Tobit regression with random effects for individual production and internet usage with respect to demographic variables.

Dependent variable:	Individual Production	Internet usage
Intercept	-3.228	7.014
Treatment *I*	9.071***	-18.150***
Treatment *TP*	9.411***	-18.083***
Arithmetic skills	5.891**	1.340
Gender	0.073	0.929
Age	0.684	-2.115
Working experience	4.165	-0.648
No. of observations	n = 111	n = 111
Log likelihood	-393.8 (Prob>χ^2)=0.005	-1699.8 (Prob>χ^2)<0.001

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

We also studied treatment effects for low- and high- ability subjects separately. We show that for low-ability subjects treatment effects were associated with a decrease in Internet usage rather than an increase in production. For high-ability subjects, treatment effects were associated with both a decrease in Internet usage and an increase in production (see Tables IV.3 and IV.4).

TABLE IV.3. Tobit regression with random effects for individual production across ability levels.

	Low-ability subjects	High-ability subjects
Intercept	7.283	-11.943
Treatment *I*	4.704	12.398**
Treatment *TP*	5.149	10.794**
Gender	0.645	-0.058
Age	0.243	1.374
Working experience	1.972	5.485
Number of observations	n = 55	n = 55
Log likelihood	-185.8 (Prob>χ^2)=0.872	-205.8 (Prob>χ^2)=0.019

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

TABLE IV.4. Tobit regression with random effects for internet usage across ability levels.

	Low-ability subjects	High-ability subjects
Intercept	-24.482	49.208*
Treatment *I*	-8.969	-27.436***
Treatment *TP*	-13.061**	-21.175***
Gender	0.747	-4.664
Age	2.169**	-0.884
Working experience	1.320	-0.143
Number of observations	n = 56	n = 56
Log likelihood	-824.2 (Prob>χ^2)=0.080	-871.6 (Prob>χ^2)<0.001

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

References

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

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

Economic Science Institute Working Papers

2012

12-31 Thomas, C. An Alternating-Offers Model of Multilateral Negotiations.

12-30 Mago, S., Sheremeta, R. and Yates, A. Best-of-Three Contest Experiments: Strategic versus psychological momentum.

12-29 Bigoni, M., Camera, G. and Casari, M. Strategies of Cooperation and Punishment among Students and Clerical Workers.

12-28 Camera, G. and Kim, J. Buyer's Equilibrium with Capacity Constraints and Restricted Mobility: A recursive approach.

12-27 Camera, G., Casari, M., and Bigoni, M. Binding Promises and Cooperation Among Strangers.

12-26 Schniter, E., Shields, T. and Dickhaut, J. Ageism & Cooperation.

12-25 Gjerstad, S. and Smith, V. Balance Sheet Crises: Causes, Consequences and Responses.

12-24 Gómez-Miñambres, J., Corgnet, B. and Hernán-Gonzalez, R. Goal Setting and Monetary Incentives: When Large Stakes Are Not Enough.

12-23 Clots-Figueras, I., Hernán González, R., and Kujal, P. Asymmetry and Deception in the Investment Game.

12-22 Dechenaux, E., Kovenock, D. and Sheremeta, R. A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments.

12-21 Rubin, J. and Sheremeta, R. Principal-Agent Settings with Random Shocks.

12-20 Gómez-Miñambres, J. and Schniter, E. Menu-Dependent Emotions and Self-Control.

12-19 Schniter, E., Sheremeta, R., and Sznycer, D. Building and Rebuilding Trust with Promises and Apologies.

12-18 Shields, T. and Xin, B. Higher-order Beliefs in Simple Trading Models.

12-17 Pfeiffer, G. and Shields, T. Performance-Based Compensation and Firm Value: Experimental evidence.

12-16 Kimbrough, E. and Sheremeta, R. Why Can’t We Be Friends? Entitlements, bargaining, and conflict.

12-15 Mago, S., Savikhin, A., and Sheremeta, R. Facing Your Opponents: Social identification and information feedback in contests.

12-14 McCarter, M., Kopelman, S., Turk, T. and Ybarra, C. Too Many Cooks Spoil the Broth: How the tragedy of the anticommons emerges in organizations.

12-13 Chowdhury, S., Sheremeta, R. and Turocy, T. Overdissipation and Convergence in Rent-seeking Experiments: Cost structure and prize allocation rules.

12-12 Bodsky, R., Donato, D., James, K. and Porter, D. Experimental Evidence on the Properties of the California's Cap and Trade Price Containment Reserve.

12-11 Branas-Garza, P., Espin, A. and Exadaktylos, F. Students, Volunteers and Subjects: Experiments on social preferences.

12-10 Klose, B. and Kovenock, D. Extremism Drives Out Moderation.

12-09 Buchanan, J. and Wilson, B. An Experiment on Protecting Intellectual Property.

12-08 Buchanan, J., Gjerstad, S. and Porter, D. Information Effects in Multi-Unit Dutch Auctions.

12-07 Price, C. and Sheremeta, R. Endowment Origin, Demographic Effects and Individual Preferences in Contests.

12-06 Magoa, S. and Sheremeta, R. Multi-Battle Contests: An experimental study.

12-05 Sheremeta, R. and Shields, T. Do Liars Believe? Beliefs and Other-Regarding Preferences in Sender-Receiver Games.

12-04 Sheremeta, R., Masters, W. and Cason. T. Winner-Take-All and Proportional-Prize Contests: Theory and experimental results.

12-03 Buchanan, J., Gjerstad, S. and Smith, V. There's No Place Like Home.

12-02 Corgnet, B. and Rodriguez-Lara, I. Are you a Good Employee or Simply a Good Guy? Influence Costs and Contract Design.

12-01 Kimbrough, E. and Sheremeta, R. Side-Payments and the Costs of Conflict.

2011

11-20 Cason, T., Savikhin, A. and Sheremeta, R. Behavioral Spillovers in Coordination Games.

11-19 Munro, D. and Rassenti, S. Combinatorial Clock Auctions: Price direction and performance.

11-18 Schniter, E., Sheremeta, R., and Sznycer, D. Restoring Damaged Trust with Promises, Atonement and Apology.

11-17 Brañas-Garza, P., and Proestakis, A. Self-discrimination: A field experiment on obesity.

11-16 Brañas-Garza, P., Bucheli, M., Paz Espinosa, M., and García-Muñoz, T. Moral Cleansing and Moral Licenses: Experimental evidence.

11-15 Caginalp, G., Porter, D., and Hao, L. Asset Market Reactions to News: An experimental study.

11-14 Benito, J., Branas-Garz, P., Penelope Hernandez, P., and Sanchis Llopis, J. Strategic Behavior in Schelling Dynamics: A new result and experimental evidence.

11-13 Chui, M., Porter, D., Rassenti, S. and Smith, V. The Effect of Bidding Information in Ascending Auctions.

11-12 Schniter, E., Sheremeta, R. and Shields, T. Conflicted Minds: Recalibrational emotions following trust-based interaction.

11-11 Pedro Rey-Biel, P., Sheremeta, R. and Uler, N. (Bad) Luck or (Lack of) Effort?: Comparing social sharing norms between US and Europe.

11-10 Deck, C., Porter, D., and Smith, V. Double Bubbles in Assets Markets with Multiple Generations.

11-09 Kimbrough, E., Sheremeta, R., and Shields, T. Resolving Conflicts by a Random Device.

11-08 Brañas-Garza, P., García-Muñoz, T., and Hernan, R. Cognitive effort in the Beauty Contest Game.

11-07 Grether, D., Porter, D., and Shum, M. Intimidation or Impatience? Jump Bidding in On-line Ascending Automobile Auctions.

11-06 Rietz, T., Schniter, E., Sheremeta, R., and Shields, T. Trust, Reciprocity and Rules.

11-05 Corgnet, B., Hernan-Gonzalez, R., and Rassenti, S. Real Effort, Real Leisure and Real-time Supervision: Incentives and peer pressure in virtual organizations.

11-04 Corgnet, B. and Hernán-González R. Don't Ask Me If You Will Not Listen: The dilemma of participative decision making.

11-03 Rietz, T., Sheremeta, R., Shields, T., and Smith, V. Transparency, Efficiency and the Distribution of Economic Welfare in Pass-Through Investment Trust Games.

11-02 Corgnet, B., Kujal, P. and Porter, D. The Effect of Reliability, Content and Timing of Public Announcements on Asset Trading Behavior.

11-01 Corgnet, B., Kujal, P. and Porter, D. Reaction to Public Information in Markets: How much does ambiguity matter?

2010

10-23 Sheremeta, R. Perfect-Substitutes, Best-Shot, and Weakest-Link Contests between Groups.

10-22 Mago, S., Sheremeta, R., and Yates, A. Best-of-Three Contests: Experimental evidence.

10-21 Kimbrough, E. and Sheremeta, R. Make Him an Offer He Can't Refuse: Avoiding conflicts through side payments.

10-20 Savikhim, A. and Sheremeta, R. Visibility of Contributions and Cost of Inflation: An experiment on public goods.

10-19 Sheremeta, R. and Shields, T. Do Investors Trust or Simply Gamble?

10-18 Deck, C. and Sheremeta, R. Fight or Flight? Defending Against Sequential Attacks in the Game of Siege.

10-17 Deck, C., Lin, S. and Porter, D. Affecting Policy by Manipulating Prediction Markets: Experimental evidence.

10-16 Deck, C. and Kimbrough, E. Can Markets Save Lives? An Experimental Investigation of a Market for Organ Donations.

10-15 Deck, C., Lee, J. and Reyes, J. Personality and the Consistency of Risk Taking Behavior: Experimental evidence.

10-14 Deck, C. and Nikiforakis, N. Perfect and Imperfect Real-Time Monitoring in a Minimum-Effort Game.

10-13 Deck, C. and Gu, J. Price Increasing Competition? Experimental Evidence.

10-12 Kovenock, D., Roberson, B., and Sheremeta, R. The Attack and Defense of Weakest-Link Networks.

10-11 Wilson, B., Jaworski, T., Schurter, K. and Smyth, A. An Experimental Economic History of Whalers' Rules of Capture.

10-10 DeScioli, P. and Wilson, B. Mine and Thine: The territorial foundations of human property.

10-09 Cason, T., Masters, W. and Sheremeta, R. Entry into Winner-Take-All and Proportional-Prize Contests: An experimental study.

10-08 Savikhin, A. and Sheremeta, R. Simultaneous Decision-Making in Competitive and Cooperative Environments.

10-07 Chowdhury, S. and Sheremeta, R. A generalized Tullock contest.

10-06 Chowdhury, S. and Sheremeta, R. The Equivalence of Contests.

10-05 Shields, T. Do Analysts Tell the Truth? Do Shareholders Listen? An Experimental Study of Analysts' Forecasts and Shareholder Reaction.

10-04 Lin, S. and Rassenti, S. Are Under- and Over-reaction the Same Matter? A Price Inertia based Account.

10-03 Lin, S. Gradual Information Diffusion and Asset Price Momentum.

10-02 Gjerstad, S. and Smith, V. Household Expenditure Cycles and Economic Cycles, 1920 – 2010.

10-01 Dickhaut, J., Lin, S., Porter, D. and Smith, V. Durability, Re-trading and Market Performance.

2009

09-11 Hazlett, T., Porter, D., and Smith, V. Radio Spectrum and the Disruptive Clarity OF Ronald Coase.

09-10 Sheremeta, R. Expenditures and Information Disclosure in Two-Stage Political Contests.

09-09 Sheremeta, R. and Zhang, J. Can Groups Solve the Problem of Over-Bidding in Contests?

09-08 Sheremeta, R. and Zhang, J. Multi-Level Trust Game with "Insider" Communication.

09-07 Price, C. and Sheremeta, R. Endowment Effects in Contests.

09-06 Cason, T., Savikhin, A. and Sheremeta, R. Cooperation Spillovers in Coordination Games.

09-05 Sheremeta, R. Contest Design: An experimental investigation.

09-04 Sheremeta, R. Experimental Comparison of Multi-Stage and One-Stage Contests.

09-03 Smith, A., Skarbek, D., and Wilson, B. Anarchy, Groups, and Conflict: An experiment on the emergence of protective associations.

09-02 Jaworski, T. and Wilson, B. Go West Young Man: Self-selection and endogenous property rights.

09-01 Gjerstad, S. Housing Market Price Tier Movements in an Expansion and Collapse.

2008

08-09 Dickhaut, J., Houser, D., Aimone, J., Tila, D. and Johnson, C. High Stakes Behavior with Low Payoffs: Inducing preferences with Holt-Laury gambles.

08-08 Stecher, J., Shields, T. and Dickhaut, J. Generating Ambiguity in the Laboratory.

08-07 Stecher, J., Lunawat, R., Pronin, K. and Dickhaut, J. Decision Making and Trade without Probabilities.

08-06 Dickhaut, J., Lungu, O., Smith, V., Xin, B. and Rustichini, A. A Neuronal Mechanism of Choice.

08-05 Anctil, R., Dickhaut, J., Johnson, K., and Kanodia, C. Does Information Transparency Decrease Coordination Failure?

08-04 Tila, D. and Porter, D. Group Prediction in Information Markets With and Without Trading Information and Price Manipulation Incentives.

08-03 Thomas, C. and Wilson, B. Horizontal Product Differentiation in Auctions and Multilateral Negotiations.

08-02 Oprea, R., Wilson, B. and Zillante, A. War of Attrition: Evidence from a laboratory experiment on market exit.

08-01 Oprea, R., Porter, D., Hibbert, C., Hanson, R. and Tila, D. Can Manipulators Mislead Prediction Market Observers?