Why Work When You Can Shirk?: Worker Productivity in an Experimental Setting

Russell Engel
Sacred Heart University

Employee shirking has the potential to be extremely costly to firms. To counter the productivity loss caused by shirking, firms may institute various incentive schemes. Previous experimental research has shown that while monitoring does decrease shirking, some subjects work without explicit financial incentives. This paper presents the experimental results of an economic experiment designed to investigate the effect of various incentive schemes on subject behavior. Subjects are allowed to engage two tasks; one task mimics work for an employer, the other task allows for gains due to shirking. We find that subjects who are given incentives to shirk do in fact shirk, but monitoring and an attainable quota lead to increased productivity. However, when the quota is unattainable, subjects revolt and engage in a high amount of shirking.

INTRODUCTION

There is a long standing and mature theoretical literature dealing with the principal-agent problem (Mirrlees (1975), Grossman and Hart (1983)). A key aspect of the principal-agent problem is the motivation of workers who are rational cheaters. Nagin (2002) defines a rational cheater as someone who will shirk when the marginal benefit of doing so exceeds the marginal cost. A rational cheater will exert low effort on the job if he thinks he can get away with it. Recent reports estimate that shirking workers cost employers billions of dollars in productivity losses yearly[1]. Employers (principals) who are aware of the financial incentives they are giving employees (agents) may introduce a monitoring system with performance goals to alleviate the perceived problem. These issues with worker motivation are difficult and complex. Consider a recent New York City court case:

On March 9, 2006, John B. Sooner, a New York City administrative law judge, recommended that Toquir Choudhri, a 14-year veteran of the city Department of Education, receive only a reprimand for disobedience, even though supervisors wanted him fired for using the Internet for personal matters[2]. Spooner wrote that Choudhri credibly stated that he completed all assignments given to him by his boss and used the internet while he awaited further assignments. These statements were corroborated by the absence of proof that Choudhri was ever criticized for poor productivity or for not completing specific assignments.[3] The New York City Chancellor of Education, Joel Klein, decided to fire Choudhri anyway. Klein stated that `the penalty of termination is appropriate and not shocking to one's sense of fairness, ... Choudhri's abuse of the Internet at the time he is supposed to be performing his job demonstrates his disinterest in the job.'[4]

The worker in the above case was fired for shirking on the job when his employer found him surfing the internet. The worker did not think he deserved to be fired because he had completed all of his assignments. The worker thought he was being monitored in regard to fulfillment of some quota, and he had fulfilled his quota, but the employer disagreed. This case demonstrates the problems caused when the monitoring system is not well delineated, but it also shows how concerned some employers are about any behavior consistent with shirking. In this paper, we will examine how well these types of schemes work when they are clearly stated and consistently enforced. We will contrast our results with previous research which showed that workers may work harder than required by explicit financial incentives.

Past experimental research (Cadsby et al. (2007), Dickinson and Villeval (2008)) has shown that some workers in a laboratory setting work without incentives. It is important to investigate if these laboratory results indicate behaviors we would see in a real work setting, or if the possibility exists that these observations are an artifact of the experimental design. This paper attempts to place subjects in a more refined laboratory setting to get a cleaner look at subject behavior towards work effort with low financial incentives. Cadsby et al. (2007), and Dickinson and Villeval (2008) were not specifically looking at effort with low incentives, so there is much remaining value to their research even if the observations of subjects working without incentives are an artifact of their design.

A common approach used in the laboratory to investigate the principal-agent problem is to give subjects a cost function and have them choose some effort level. Nalbantian and Schotter (1997), for example presents an experiment in which a subject is monitored with probability. If the subject is not expending a certain level of `effort', he will be terminated. The `effort' in this case is not physical exertion but rather a figurative effort. This number they pick is indeed interpreted as effort and therefore has the property that effort is now explicit. While this matches clearly with their models, it is not clear that subjects perceive this choice as analogous to physical or mental exertion. The (e) chosen by the subject is costly to the subject but it is possible that this is too abstract to model real work. Putting the workers through a real effort experiment will allow an answer as to whether simply choosing effort garners the same behavior as exerting effort, and if it does not, one can be confident that the real effort experiment is a better proxy for the workplace.

While an agent's outside option can be represented rather easily in a theoretical framework, it is not so trivial to do in a laboratory. To think of this one must place himself in the position of the subject in the experiment. The subject arrives at the laboratory and is assigned a computer terminal. They are given the option of engaging some task and earning X or he can take the outside option (sit still) and earn Y. If the disparity between the outside option and the participation option is not large, there exists the possibility that the subject will engage the game to avoid boredom.

The aforementioned research by Cadsby et al. (2007), and Dickinson and Villeval (2008) found that subjects contribute effort even when they have no financial reason to do so. The idea of a moral imperative not to shirk is given as a reason for this behavior by Dickinson. Cadsby's experiment gives subjects seven scrambled letters and the subjects are instructed to make as many words as possible in a given time period. The subjects are allowed to choose a piece rate or a flat rate scheme, the piece rate will pay them per word, and the flat rate will give them some stated amount with no requirement on word creation. These incentives should lead students who think they are endowed with word creation ability and low effort cost to choose the piece rate scheme and the subjects who have high effort cost or are not good at word creation to choose the flat rate scheme. The interesting observation is that there are significantly more than zero words created by the subjects who choose the flat rate scheme. This is contrary to the incentive structure, and one would not believe the students feel a moral imperative to unscramble letters. Dickinson and Villeval (2008) use a real effort task and they observe that some of the subjects (25%) contribute at or above the desired output level even when monitoring is set to zero. Dickinson, as noted earlier, suggests this as either intrinsic motivation or integrity and commitment to moral principles. It is possible that subjects feel a moral imperative because they are interacting with a human principal, but that is not the case in Cadsby et al. (2007). The intrinsic motivation argument is plausible in both experiments. The subjects might enjoy unscrambling letters, and they might enjoy moving along a curve to get a high value, but it is not clear that observations made under these conditions should be interpreted as analogous to intrinsic motivation one might experience in the real world. In both experiments, the subjects could engage in effort, or do nothing. It is possible that they were engaging the tasks because they were bored. This is similar to Choudri's claim that he browsed the internet only because he had no other work to do. If the subjects in the experiment do not engage the task, they have nothing else to do.

Other experiments have shown that outside options have important effects. Lei, Plott, and Noussair (2001) show that excess trading in an asset market can be reduced by giving subjects something to do beyond trading in the asset market. Pevnitskaya and Palfrey (2008) show that over entry into an auction can be reduced by allowing subjects an outside option of a computerized version of rock-paper-scissors. Van Dijk et al. (2001) conduct an experiment where subjects are enabled to work on two tasks in the same period. The earnings from one task go in to a group account while the earnings from the other task go in to a private account, similar to a public goods game. This alleviates the effort only due to boredom problem that is plausible in the Cadsby et al. (2007) and Dickinson (2004) papers. The specific task the paper uses has the subjects search a grid looking for the highest payoff. The idea of having two of these for the subjects to play cures the boredom critique.

The experimental design for this paper allows for subjects to play a valuable outside option. This should eliminate play in the primary task when there is no financial incentive to play the primary task. Then the question of how much monitoring is necessary to get the desired amount of effort can be addressed. The desired amount of effort will be explicit in terms of a quota.

The paper proceeds as follows: Section 2 contains the experimental design, section 3 contains a simple model of predicted behavior in the experiment, and section 4 reports analysis of the results of the experiment. Finally, there will be some concluding remarks.

EXPERIMENTAL DESIGN

The experiment conducted for this paper was designed to give subjects incentives similar to those faced by many workers. There are 30 periods in the experiment, and each period lasts 45 seconds with 30 seconds between each period.. The experiment consists of a primary task (Task A) and a secondary task (Task B). The subjects can choose to split effort, at their discretion, between the two tasks. Task A is designed to mimic work for an employer. The subject will view a randomly generated four-digit number and his task is to type the same number in the space provided. Every time a subject has completed typing the number, he can hit a button and start on a new randomly generated number. An example of this can be seen in Table 1. The payoff to the subjects depends on the quota they face and a monitoring level. The subjects earn an effective wage of 300 ECUs per period. If they are monitored and the amount of four-digit numbers typed is less than the quota they are considered fired without pay, and their earnings are zero. If a subject is not monitored, he earns 300 ECUs from Task A regardless of whether or not his quota is met.

TABLE 2
TASK A

Probability of being monitored:	25%
Quota:	2
Number Completed:	2
Type this:	7072
Enter here:	
You entered:	

The other task subjects can engage in (Task B) is a matching pennies game, shown in Table 2, that they will play against a computer opponent. The computer is playing a Mixed-Strategy Nash Equilibrium of choosing heads 50% of the time and tails 50% of the time. The matching pennies game was chosen because it is a cognitively easy task, and the chosen payoffs have the property that the expected utility is increasing in the amount of time spent on the task. This is intended to mimic the utility a subject would receive from shirking, whether it be reading the newspaper, talking to friends, or browsing the internet. Without Task B, it is possible that the subject would play Task A simply because they are bored and find Task A more stimulating than sitting quietly. With Task B, the subjects could still choose to sit still, but they now have the alternative action involving activity of an outside option that they can earn money by playing.

TABLE 2
MATCHING PENNIES

		Computer	
		Heads	Tails
Subject	Heads	2,0	0,2
	Tails	0,2	2,0

Any earnings from Task B are in addition to the 300 ECUs earned in Task A, but they are not guaranteed. If the subject is monitored in task 1, and their number completed is less than their quota, any earnings they made in Task B are wiped out. This is a punishment akin to getting fired. One could argue that the subject should keep their outside earnings since an employer can fire an employee without pay, but cannot take away the utility they received from shirking. There will be some fine associated with getting fired and for simplicity the amount of that fine will be equal to the earnings form Task B. This is done to exact some punishment beyond lost wages. In the laboratory, the worker will be rehired the next period, so just taking away the wages for one period is not punitive enough. Regardless of how much is potentially earned on Task B, a subject who is monitored and has not met his quota always earns 0 ECUs for the period.

In order to observe the subjects responses to a variety of situations, there were 6 different monitoring levels (0%, 15%, 25%, 50%, 75%, and 100%) and 5 different quotas (8, 12, 15, 18[5], and 25). This lead to 30 combinations of monitoring levels and quotas and each subject faced each combination once. This resulted in the 30 periods mentioned before. The order that the subjects saw the various quotas and monitoring levels was determined randomly prior to the experiment. At the end of every period the subjects would see a screen that would inform them whether or not they were monitored, and their payoff at the end of each period.

The experiment was conducted over four sessions and 32 subjects participated. The subjects for the experiment were undergraduate students at the Florida State University. The experiment was computer based and conducted with z-Tree software (Fischbacher 1999). When the subjects arrived, they were assigned to computer terminals. The instructions for the experiment (see Appendix) were read aloud and subjects had a chance to ask questions. The subjects were paid a $10.00 show up fee, and were able to earn money based upon their performance in the experiment. The average earnings per subject for the experiment were $24.96.

THEORY

Given the incentive structure in the experiment, we can now develop a model to predict the behavior of the subjects. Let t be the amount of time a subject spends on Task A, let $m \in (0,1)$ be the monitoring level the subject faces at the beginning of each period, let $Q \in (3,25)$ be the quota the subject faces at the beginning of each period. Let $\phi_i(t)$ be the payoff a subject receives for working on Task B ($\frac{d\phi_i}{dt} < 0$ because the more time a subject spends on Task A, the less time they have for Task B). The probability that a subject meets the quota (Q) in time (t) is represented by the term $P_i(t,Q)$. This probability is increasing in t $\left(\frac{dP_i}{dt} > 0\right)$, because the more time spent on Task A, the more likely they are to meet the quota; $P_i(t,Q)$ is decreasing in Q $\left(\frac{dP_i}{dQ} < 0\right)$ because the higher the quota, the more difficult it is to meet it given the 45 second time constraint. Both $\phi_i(t)$, and $P_i(t,Q)$ are individual specific due to the heterogeneity in subject ability (subjects who are good at Task A could also be good at Task B, but that is irrelevant to the analysis of this problem). Recall that subjects can earn 300 ECUs each period if they meet the quota, or are not monitored. The payoff function subjects face can be represented as:

$$\Pi_i(t,Q,m) = mP_i(t,Q)(300 + \phi_i(t)) + (1-m)(300 + \phi_i(t)) \quad (1)$$

This payoff function shows that if a subject is not monitored, they receive $(300 + \phi_i(t))$; if they are monitored they only receive $(300 + \phi_i(t))$ if they meet the quota, which happens with probability $P_i(t,Q)$. In order to demonstrate some implications of this basic model we will assume simplifying assumptions for ϕ_i and P_i. Let $\phi_i(t) = a - bt$, and

$$P_i(t,Q) = \begin{cases} \frac{ct}{Q} & for \quad ct \le Q \\ 1 & for \quad ct > Q \end{cases}$$

The results we derive are not special to these functional forms, but will hold for a broad range of functions satisfying the conditions: $\frac{d\phi_i}{dt} < 0$ and $\frac{dP_i}{dt} > 0$. Equation 1 now becomes:

$$\Pi_i(t,Q,m) = m\frac{ct}{Q}(300 + a - bt) + (1-m)(300 + a - bt) \tag{2}$$

The general problem is to maximize Π_i given some $t \le T$.

$$\max_{t \le T} \Pi_i(t,Q,m) = mP_i(t,Q)(300 + \phi_i(t)) + (1-m)(300 + \phi_i(t)) \tag{3}$$

The solution is:

$$t^* = \frac{1}{2}\frac{300mc + mca - bQ + mbQ}{mcb}. \tag{4}$$

Given this solution we can show the following points.

Implication 1: *When the monitoring level is zero, a subject should not work on Task A.* When there is no monitoring, equation 1 becomes:

$$\Pi_i = (300 + \phi_i(t)) \tag{5}$$

As stated earlier, the function $\phi_i(t)$ is declining everywhere in t, any time spent on Task A means less earnings can be made in Task B. So, $\frac{d\Pi_i}{dt} < 0 \forall t$. Therefore, to max $U_i = (300 + \phi_i(t))$, a subject should choose $t = 0$, and we should observe subjects playing only Task B.

Implication 2: *As the monitoring level increases, a subject should spend more time on Task A.*

$$\frac{dt^*}{dm} = \frac{d\left(\frac{1}{2}\frac{300mc + mca - bQ + mbQ}{mcb}\right)}{dm} = \frac{1}{2}\frac{Q}{m^2c} > 0 \tag{6}$$

As the monitoring level increases, shirking behavior is more likely to be punished. To avoid this punishment, a subject has to meet the quota. So, as the monitoring level increase, a subject should spend more time on Task A.

Implication 3: *If the subject chooses to work on Task A, once he meets the quota, he should spend no additional time on Task A.*

If $\exists\ \hat{t}\quad s.t.\ P_i\left(\hat{t},Q\right)=1$, any $t^*>\hat{t}$ will reduce the subjects expected utility. Let $P_i\left(\hat{t},Q\right)=1$, and let $t^*>\hat{t}$.

If subject chooses $t=\hat{t}$: $\Pi_i = m\left(300+\varphi_i(\hat{t})\right)+(1-m)\left(300+\varphi_i(\hat{t})\right)=300+\varphi_i(\hat{t})$.

If subject chooses $t=t^*$: $\Pi_i = m\left(300+\phi_i(t^*)\right)+(1-m)\left(300+\phi_i(t^*)\right)=300+\phi_i(t^*)$.

$300+\varphi_i(\hat{t})>300+\varphi_i(t^*)$ since $t^*>\hat{t}$ and ϕ_i is decreasing in t.

Implication 4: *For a given monitoring level, the subject will increase observed effort as quota increases until some point, as the quota gets relatively large, they will then exert less effort on Task A.*

Given that:

$$P_i(t,Q)=\begin{cases}\dfrac{ct}{Q} & for \quad ct\le Q\\ 1 & for \quad ct>Q\end{cases}$$

we have to consider that a subject should never choose

$$t^*=\frac{1}{2}\frac{300mc+mca-bQ+mbQ}{mcb}>\hat{t}$$

where $\hat{t}$ is defined by $P_i\left(\hat{t},Q\right)=1$. In our example $P_i(t,Q)=1$ when $ct=Q$. Therefore $\hat{t}=\dfrac{Q}{c}$.

The optimal amount of effort a subject will choose is

$$\min(t^*,\hat{t})=\min\left(\frac{1}{2}\frac{300mc+mca-bQ+mbQ}{mcb},\frac{Q}{c}\right).$$

$$t^*<\hat{t}\text{ when }Q>\frac{300+a}{b(m+1)}mc$$

$$\frac{d\hat{t}}{dQ}=\frac{d\left(\dfrac{Q}{c}\right)}{dQ}=\frac{1}{c}>0 \tag{7}$$

$$\frac{dt^*}{dQ}=\frac{d\left(\dfrac{1}{2}\dfrac{300mc+mca-bQ+mbQ}{mcb}\right)}{dQ}=\frac{-1+m}{mc}<0 \tag{8}$$

We see in Equation 7 that for $Q > \frac{300+a}{b(m+1)}mc$ a subject's effort will be increasing with Q, and in Equation 8 that for $Q < \frac{300+a}{b(m+1)}mc$ a subjects effort will be decreasing in Q. Fundamentally what this is saying is that subjects will maximize their probability of earning the 300 ECU wage when the quota is relatively low, but at some point increasing the quota will lead to a reduction in the effort (amount of numbers typed) given to Task A.

Implication 5: *If the subject cannot meet the quota even if they were to spend all of their time on Task A, he should not spend any time on Task A.*

If $P(t,Q) = 0 \quad \forall t \le T$, we see that equation 1 reduces to $U_i = (1-m)(300+\phi_i(t))$. Again, the subject should only work on Task B.

Implication 6: *If the subject cannot meet the quota even if they were to spend all of their time on Task A, and the monitoring level is 100%, observations of the subjects playing either task would imply that they would rather engage the experiment than sit still.*

If $m = 1$, equation 1 becomes:

$$\Pi_i = P_i(t,Q)(300+\phi_i(t)) \tag{9}$$

If $P(t,Q) = 0 \forall t \le T$, then equation 9 is equal to zero. The subject gets zero financial gain from either task, so the subject should not engage either task unless they prefer it to sitting still.

Given these implications, we should expect the subjects to only play Task B when the monitoring level is zero, and to increase their effort on Task A as monitoring is increased. The subjects should also increase their effort on Task A when the quota increases if the quota is relatively low. An increase in the quota beyond some point will lead the subjects to reduce effort on Task A. If a subject meets the quota in a given period, he should then switch to playing only Task B as he longer has any financial incentive to play Task A. When the quota is unattainable, the subjects should not try to reach the quota but instead spend all of their time on Task B. If the quota is unattainable and the subject is definitely going to be monitored, any engagement of either Task A or Task B implies that the subject prefers engaging the experiment to sitting still.

EXPERIMENTAL RESULTS

Result 1: *When the monitoring level is zero, ninety-one percent of the subjects do not meet or exceed the quota. However, when the quota is small, many subjects meet that quota.*

Table 3 shows that subjects are highly unlikely to exert effort when they are not being monitored. Whereas *Implication 1* suggests that subjects should exert zero effort when the monitoring level is set to zero, the data show that some subjects do meet the quota when the monitoring level is 0%. Dickinson and Villeval (2008) found that 25% of subjects contributed at or above the desired output level; we find less than that. Table 3 shows that there are 14 out of 152 (09%) observations where the subject meets or exceeds the quota. But, we have to note that not all subjects could reach all quotas. Only one subject in the entire experiment met $Q = 25$, five subjects met $Q = 18$, seventeen subjects met $Q = 15$, all thirty two subjects met $Q = 8$, and

$Q = 12$. So, to get another estimate of worker effort when monitoring level is zero we can use the the highest observed effort level for each subject to note if someone had the ability to meet a given quota. This leads to 14 out of 87 (16%) workers meeting or exceeding the quota when the monitoring level is zero. This estimate is not perfect as it has the ability to underestimate shirking. Some subjects may have been able to meet quotas that they were not observed to meet if they had worked harder. Best stated, when monitoring level is zero, the percentage of workers that meet or exceed the quota is in the range of 9% to 16%. Furthermore, most of the observations where workers meet or exceed the quota, 10 out of 14 (71%), occur when the quota is relatively low (Q=8). When $Q > 8$ and monitoring is zero, only 4 out of 55 (07%) subjects meet or exceed the quota.

TABLE 3
SUBJECT EFFORT WHEN MONITORING LEVEL IS ZERO

Number of four-digit numbers typed when monitoring level = 0%									
		0-5	6-10	11-15	16-20	21-25	NC=Q	NC>Q	Period
Quota	8	21	8	3	0	0	7	3	7
	12	26	2	3	1	0	2	1	10
	15	29	2	1	0	0	1	0	17
	18	21	1	2	0	0	0	0	21
	25	14	13	5	0	0	0	0	2

To test *Implications 2 and 4*, we run a linear random effects panel regression. This type of regression controls for omitted variables that differ between subjects and omitted variables that vary within subjects over time. We will regress the amount of four-digit numbers typed on monitoring level (*Implication 2*), quota less than 15, and quota greater than 15 (*Implication 4)*. The regressor, $dummy_{Q<15}$, is set to 1 when $Q < 15$ and zero when $Q > 15$, $dummy_{Q>15}$ is set to 1 when $Q > 15$ and zero when $Q < 15$. If subjects respond in the experiment as predicted by *Implication 4* they will increase effort up until some Q, and then begin to reduce effort. If this is the case, we need a constant for both situations. The regressor $dummy_{period \leq 5}$ is set to 1 for periods 1-5 and zero otherwise. This is used to determine if behavior in the first five periods is significantly different than play in latter periods. This could be a factor if the subjects experience any learning effects in the first five periods. Prior to the experiment, it is unlikely that the subjects know how many four-digit-numbers they can type in 45 seconds; the subjects may use the initial periods to gauge their ability. Table 4 reports the results of this regression.

Result 2: *Monitoring has a positive significant impact on effort given to task 1.*

Table 4 shows that as monitoring increases subjects choose to exert more effort (type more numbers) on the monitored task. This result is consistent with *Implication 2*. Specifically, if monitoring goes up by one percentage point, the average subject increases the amount of four-digit numbers by 7.9%. Figure 1 is a scatterplot of four-digit-numbers typed on quota for a given monitoring level. Figure 1 allows us to see the impact of monitoring in a sequence. There are 160 observations in each cell. The data have been jittered due to the high level of overlap of the data points. The dark spots in each cell represent the most overlap. We see that when the monitoring level is low, observations are clustered around zero effort. As monitoring increases the mass of observations starts to move up, and when monitoring is at 100% we see a very strong correlation between the quota and the amount typed. If we look at Figure 1 we see that the subjects are exerting effort where monitoring is zero and $Q = 25$. The subjects saw this combination of quota and monitoring level early in the experiment (period 2). The significance of the regressor $dummy_{period \leq 5}$ in Table 4 shows that subjects were playing differently early in the experiment. It is possible that the subjects were exploring the game space.

TABLE 4
RESULTS OF LINEAR RANDOM EFFECTS PANEL REGRESSION OF NUMBER OF FOUR-DIGIT NUMBERS TYPED ON POSSIBLE EXPLANATORY VARIABLES.

	Coef.	Std. Err	P-value
Monitoring*	.0786	.0040	0.000
quota(<15)*	.5140	.0742	0.000
quota(>15)*	-.9551	.0440	0.000
$dummy_{Q<15}$	1.404	.8858	0.281
$dummy_{Q>15}$*	6.257	1.002	0.000
$dummy_{period<5}$*	1.850	.3807	0.000
Number of Obs	960		

Result 3: *Once subjects meet the quota in Task A, most switch to Task B exclusively.*

If we refer back to Figure 2 we see few observations where the amount typed is greater than the quota. Consider the case where monitoring is equal to 100%, there are 70 instances where the amount typed is greater than or equal to the quota, of these 70 instances, 62 hit the quota exactly. In Figure 2, we can observe that the number typed is rarely above the quota.

Result 4: *The quota level (for quota less than 15) has a positive significant impact on effort. The quota level (for quota greater than 15) has a negative significant impact.*

Table 4 shows that, consistent with *Implication 4,* when the quota is relatively small subjects increase their effort on Task A. However, when the quota is greater than 15 subjects' effort is decreasing as the quota increases. We can see in Figure 2 that when the quota increases from 8 to 12, there is an increase in the amount of effort given to Task A. When the quota increases to 15, we can see a drop off in the effort given to Task A, particularly when monitoring is less than 50%.

FIGURE 1
SUBJECT EFFORT WHEN MONITORING LEVEL IS ZERO

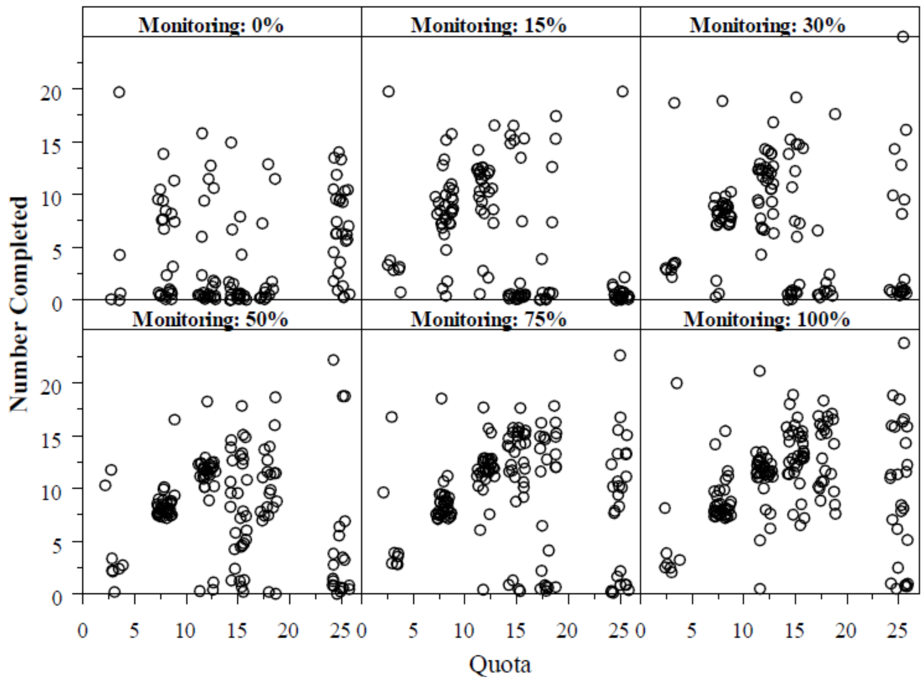
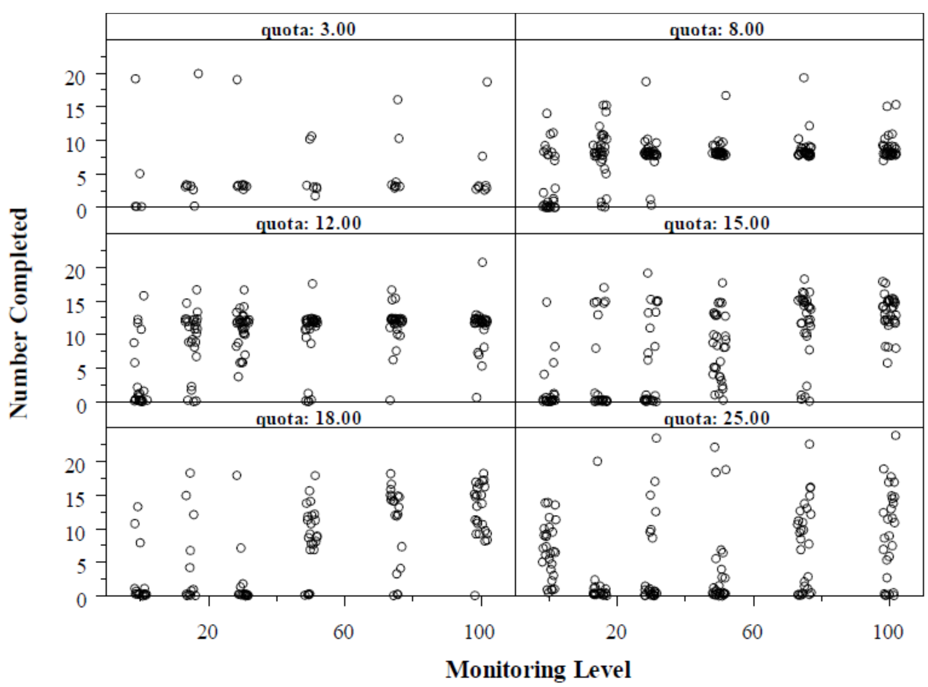


FIGURE 2
SUBJECT EFFORT GIVEN A CONSTANT QUOTA



Result 5: *When the quota is set to 25, this is effectively* $P(t,Q) = 0 \quad \forall t$, *and monitoring is less than 100%, subjects prefer to spend more time on Task B than Task A.*

Prior to the experiment, the quota level of 25 was believed to be so high that no subject could hit it, and would therefore be implicitly unattainable. This held true for 30 of the 32 subjects, but one subject was able to reach $Q = 25$ and another was able to type 24 four-digit numbers. In the minds of these two subjects, the quota of 25, was probably not unattainable. The other subjects never typed as much as 20 four-digit numbers. So, for 30 of the 32 subjects, $Q = 25$ should have been inferred as unattainable. *Implication 3* demonstrated that if a subject could not meet the quota even if they spent all of their time on Task A, then they should not try to meet the quota. Table 5 shows that when subjects think they cannot meet the quota, they will not try to meet the quota. If we contrast this unattainable quota with the high quota of 18, we see that the subjects exert much more effort when there is at least some possibility that they might reach the quota. Consider, for example, the case where monitoring = 50%.When the quota is 25, the number of subjects that type five or fewer four-digit numbers is 26 out of 32 (81%); when the quota is 18, the number of subjects that type five or fewer four-digit numbers is 5 out of 24 (21%). If we look to Figure 2 we can see that effort given to Task A is much greater when the quota is 15 compared to when the quota is 25. *Result 1* showed that some subjects will meet easily attainable quotas even without financial incentives. *Result 5* shows that unattainable quotas will cause subjects to greatly reduce effort.

TABLE 5
SUBJECT EFFORT WHEN Q=25

		number of four-digit numbers typed when Q=25						
		0-5	6-10	11-15	16-20	21-25	NC=Q	Period
M	0%	14	13	5	0	0	0	2
	15%	31	0	0	1	0	0	24
	30%	25	3	2	1	1	0	13
	50%	26	3	0	2	1	0	19
	75%	15	6	8	2	1	0	8
	100%	12	5	9	4	2	1	30

TABLE 6
SUBJECT EFFORT WHEN Q=18

		number of four-digit numbers typed when Q=18						
		0-5	6-10	11-15	16-20	21-25	NC=Q	Period
M	0%	21	1	2	0	0	0	21
	15%	20	1	2	1	0	1	20
	30%	22	1	6	1	0	1	28
	50%	5	9	8	2	0	1	6
	75%	9	1	11	3	0	1	26
	100%	1	6	10	7	0	1	23

|

Result 6: *When the Quota is set to 25 (* $P(t,Q) = 0 \quad \forall t$*), and the monitoring level is set to 100%, subjects do not sit still, they engage both Task A and Task B.*

Table 7 shows when monitoring is 100% and $Q = 25$ many subjects exert effort even though they know they will not meet the quota. One subject was extremely fast and did meet the quota, another made it to 24. No one else made it to 20. Seven subjects engaged neither task, two subjects engaged both, three subjects played only the outside option, and twenty engaged only the typing task. Similar to this result, both Cadsby et al. (2007), and Dickinson and Villeval (2008) found subjects contributing without financial incentives. *Result 6* lends support to the claim that subjects will engage tasks simply because they are bored. One could argue that since the only possible way a subject could get any earnings was to try and reach the quota, effort given to Task A was not play due to boredom. While this is plausible, it is not likely given the intensity of effort given to Task A. If we look at Tables 5 and 6, when monitoring is set to 100%, we see that work intensity is higher for $Q = 18$. Consider that when $Q = 25$ only 15 out of 32 (47%) subjects type more than 10 numbers, but when $Q = 18$ 17 out of 24 (71%) subjects type more than 10 numbers. So, the workers who choose to give effort to Task A when monitoring is 100% and $Q = 25$ do not appear to be exerting high effort, and that is consistent with play due to boredom.

TABLE 7
SUBJECT EFFORT WHEN MONITORING LEVEL=1

		number of four-digit numbers typed when m=100%							
		0-5	6-10	11-15	16-20	21-25	NC=Q	NC>Q	Period
Quota	8	0	28	4	0	0	27	4	5
	12	2	3	26	0	1	24	1	22
	15	0	4	25	3	0	9	3	14
	18	1	6	10	7	0	1	0	23
	25	12	5	9	4	2	1	0	30

CONCLUSION

This paper presented an experimental study of worker productivity. The main objective of the paper was to examine the effect of various incentive schemes on subject behavior in an environment that is meant to mimic a work setting where the worker has the ability to shirk by engaging in some task other than work for the employer. A secondary objective of the paper was to examine past experimental results that showed subjects exerting effort when no financial incentive to do so existed, and determine if this was consistent with play due to boredom.

The key finding of this paper is that workers in a laboratory work setting with an outside option available to them do shirk, but monitoring is quite successful at reducing shirking. In fact, when the quota is not difficult to attain, very little monitoring is necessary to gain subject compliance. However, when the quota is unattainable, the subjects revolt and exert very little effort. When there is no monitoring, we do observe that some subjects may still have some intrinsic motivation to play the primary task, but less than previous work reported. Another finding is that by allowing subjects to participate in an outside option, we are able to mitigate play that is due to boredom. And by implicitly disallowing earnings we are able to see that subjects prefer to play tasks instead of sitting still.

Subjects in this experiment were observed to work until they met their quota and then switched to the outside option. This paper began with the story of a worker who was fired for a lack of productivity. But it is not clear that he was shirking. Once he completed his assignments, or met his quota, he played the outside option available to him. This is the exact behavior subjects exhibited in the laboratory.

Future work will examine if subjects who are monitored over any shirking behavior exert more effort on their primary task than subjects monitored over performance goals. The issue of self selection will also be examined by paying subjects based on the amount of numbers they type as opposed to a flat wage. Subjects will then be allowed to choose the pay for performance scheme or the flat wage and monitoring scheme.

ENDNOTES

1. Frauenheim, Ed. "Stop Reading This Headline and Get Back to Work." *CNET*, Monday, July 11, 2005, http://news.com.com/Stop+reading+this+headline+and+get+back+to+work/2100-1022_3-5783552.html
2. Klopott, Freeman. "Should You Be Fired for Using the Internet While at Work?" *PC World*, Tuesday, May 02, 2006, http://pcworld.com/article/id,125597-page,1/article.html?RSS=RSS
3. Department of Education v. Choudhri, New York City Office of Administrative Trials and Hearings, No. 722/06 (3/9/06)
4. Associated Press. "NYC Fires Man For Web Surfing At Work." *CBS News*, http://www.cbsnews.com/stories/2006/05/06/tech/main1596034.shtml
5. The quota level of 18 only applies to subjects 9-32. The first 8 subjects had a quota of 3 instead of 18.

REFERENCES

Cadsby, C. Bram, Fei Song, and Francis Tapon. (2007). Sorting and Incentive Effects of Pay-for-Performance: An Experimental Investigation. *Academy of Management Journal*, 50 (6), 1554–1557.

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

Fischbacher, Urs (1999). z-Tree - Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, Working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

Grossman, S. J., and O.D.Hart (1983). An Analysis of the Principal-Agent Problem. Econometrica 51(1), pages 7-45.

Lei, V., C. Noussair, and C. Plott (2001). Non-Speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality. Econometrica 69, (4), pp. 831-859.

Mirrlees, J. A.(1975). The Theory of Moral Hazard and Unobservable Behaviour I, Mimeo, Oxford (Published in review of Economic Studies in 1999).

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

Nagin, D. et al. (2002). Monitoring, Motivation and Management: The Determinants of Opportunistic Behavior in a Field Experiment. American Economic Review, 92 (4), pp. 850-73.

Palfrey, T. and S. Pevnitskaya (2008). Endogenous Entry and Self-selection in Private Value Auctions: An Experimental Study. Journal of Economic Behavior and Organization, *2008, 66, 731-747*

van Dijk, F., Joep Sonnemans and Frans van Winden (2001). Incentive Systems in a Real Effort Experiment. European Economic Review, 45, 187-214

I would like to thank Mark Isaac, Tim Salmon, and the participants of The Florida State University Experimental Economics Readings Group for helpful comments. Funding for the experiment was provided by the John and Hallie Quinn fellowship. All remaining errors are my own.

APPENDIX

EXPERIMENT INSTRUCTIONS

In this part of the experiment you will be able to work on two tasks. You can split your time among the tasks however you choose. Meaning you can spend all of your time on task A and none on task B, all of your time on task B and none on task A, or some combination of task A and task B. Each round in this experiment is 45 seconds long.

In the box marked task A on your handout a number is displayed. Your task is to type the number you see in the box provided. Every time you click the " OK" button, a new number will come up and you can go through the task again. In each round of the experiment, the default payment for task A is 300 ECUs (1000 ECUs = $1.00). Payment will be further described below.

If you look at your handout, you will see two other pieces of information for task A (Quota and Probability of being monitored).

Monitoring means that the computer is checking to make sure that you have met the quota (defined in next paragraph). The " monitoring level" tells you how likely it is that the computer is going to check on you. If the " monitoring level" is 0%, that means the computer will not check on you. If it is 100%, that means the computer will definitely check on you.

If you are monitored: The " quota" is the minimum number of task A numbers you have to type in order to earn 300 ECUs. In other words, to earn the 300 ECUs, you must type at least the amount of numbers specified by the quota and they must be typed correctly. Your screen will update the amount of completed correct numbers you have typed (" number completed" on your handout). If you do not type enough numbers correctly to cover the quota, your earnings for the round are zero.

If you are not monitored: If you meet the quota you earn 300 ECUs. If you don't meet the quota, you earn 300 ECUs.

The monitoring level is stated at the beginning of every round. At the end of each round, the computer will generate a random number between 1 and 100 with all numbers being equally likely to determine whether or not you are in fact monitored. For example, assume you enter the round and the " monitoring level" is set to 40 (this is a 40% chance that you will be monitored). If the number drawn at the end of the round is between 41 and 100, you will not be monitored. If the number drawn is between 1 and 40, you will be monitored.

Alternatively, you may also choose to work on task B (the bottom section of the handout).

In this task you will be playing a game against the computer. The computer is programmed to pick either option H or option T. You also have the ability to pick option H or option T (you do this by choosing H or T and then clicking the " ok" button).

If both of you pick the same option you win, and earn 2 ECUs, if the choices don't match, the computer wins, and you earn zero ECUs. However, if you are monitored in task A, and you did not meet the quota, whatever earnings you had accumulated in task B will be wiped out.

The example in the handout has the monitoring level at 25 (that is a 25% chance that you will be monitored), and the quota is set at 2. If the " number completed" at the end of the 45 second round is greater than the quota, or equal to the quota, you will earn 300 ECUs from task A plus whatever you earned on task B. If " number completed" is less than the quota, two things can happen:

1. You are monitored, and you earn zero for the round (zero from task A and zero from task B)

2. You are not monitored, and you earn 300 ECUs from task A plus whatever you earned on task B.

On the second page of your handout, the example shows a case where monitoring occurred. However, since the quota was met, the earnings from task A are 300 ECUs. Your total earnings are 306, because you earned 6 ECUs from task B. If the quota had not been met, the task A earnings would be zero, and your earnings from task B would be wiped out, so your total earnings for the round would be zero.