I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10546

The Effectiveness of Incentive Schemes in the Presence of Implicit Effort Costs

Sebastian J. Goerg
Sebastian Kube
Jonas Radbruch

FEBRUARY 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10546

The Effectiveness of Incentive Schemes in the Presence of Implicit Effort Costs

Sebastian J. Goerg
Florida State University and MPI for Research on Collective Goods

Sebastian Kube
University of Bonn and IZA

Jonas Radbruch
University of Bonn and IZA

FEBRUARY 2017

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effectiveness of Incentive Schemes in the Presence of Implicit Effort Costs*

Agents' decisions to exert effort depends on the provided incentives as well as the potential costs for doing so. So far most of the attention has been on the incentive side. However, our lab experiments underline that both the incentive and cost side can be used separately to shape work performance. In our experiment, subjects work on a real-effort task. Between treatments, we vary the incentive scheme used for compensating workers. Additionally, by varying the available outside options, we explore the role of implicit costs of effort in determining workers' performance. We observe that incentive contracts and implicit costs interact in a non-trivial manner. Performance reacts significantly to changes in implicit effort costs under low-powered piece-rate and target-based bonus contracts, but not under a high piece rate contract. In addition, comparisons between the incentive schemes depend crucially on the implicit costs.

JEL Classification: C91, D01, D03, D24, J22, J24, J33, L23, M52

Keywords: workers' performance, work environments, implicit cost, opportunity costs, incentive schemes

Corresponding author:
Sebastian J. Goerg
Department of Economics
Florida State University
113 Collegiate Loop
Tallahassee, FL 32306-2180
United States

E-mail: sgoerg@fsu.edu

* We are grateful to John Hamman and Felix Schran for helpful comments on earlier drafts. We thank Anne Mertens for excellent research assistance.

1 Introduction

What are the determinants of effort provision, and how to incentivize agents to exert high effort? Most studies addressing these questions usually focus on the compensation side, investigating effort responses to fixed and variable wages (Lazear, 2000; Carpenter, 2016), fair wages (Cohn *et al.*, 2015), or other contractual details of the incentive scheme (Winter, 2004; Herweg *et al.*, 2010; Goerg *et al.*, 2010). Yet, agents' behavior also depends on additional non-monetary features of the work environment. Examples for such additional influences include task-specific intrinsic motivation (Deci, 1971), recognitions and awards (Kosfeld and Neckermann, 2011; Bradler *et al.*, 2016), personal goals (Koch and Nafziger, 2011; Goerg and Kube, 2012; Corgnet *et al.*, 2015a), and restrictions on behavior (Falk and Kosfeld, 2006). In this paper we demonstrate that the opportunity costs of effort, which crucially depend on the work environment, play a central role for effort provision in general and for the effectiveness of incentive schemes in particular.

More generally, effort provision by an agent is determined not only by the incentives provided for a given task, but also by the effort costs an agent faces. Effort costs can be financial expenditures, but more importantly they comprise opportunity costs of foregone alternative activities (see for example Holmstrom and Milgrom, 1987). Therefore, incentives to perform in a given task can generally be provided by either setting the incentive scheme or by controlling the outside options of an agent (Holmstrom and Milgrom, 1991). Whereas, the incentive side of the problem has been extensively studied, the interaction of outside activities and incentive schemes has been largely ignored. We intend to close this gap with the help of a real effort experiment in which we manipulate opportunity costs and incentives.

To manipulate the opportunity costs we implement three different work environments resulting in different implicit costs.[1] The first environment, Fix, is a standard lab environment in which subjects have to stay and perform a real effort task for a fixed period of time. In the other two environments we increase the implicit costs by giving subjects the opportunity to reduce the time they work on the task and allow them to allocate their time differently. In the environment Inet subjects can either work on the task or surf the internet; however, they have to stay in the lab for the same time as in Fix. In the environment Free, subjects are free to quit the task and leave the lab early. On the second dimension we vary the incentive schemes under which the subjects are working. We implement two different piecerate schemes (Piecerate-Low, Piecerate-High) and two non-discretionary bonus schemes. In the first bonus scheme, the necessary output thresh-

[1] Opportunity costs are the sum of the direct and explicit effort costs a worker bears as well as the implicit effort costs which constitute the foregone utility by not allocating these resources towards an alternative activity.

old is easy to achieve (Bonus-Easy) and in the second one it is nearly impossible to achieve (Bonus-Hard). As expected, we observe higher output in the Fix-environment compared to the Inet- and Free-environments with increased implicit costs. However, in the Fix-environment all four incentive schemes result in similar outputs, although marginal incentives vary substantially. In Inet- and Free-environments subjects respond to incentives and we observe positive elasticities for the output.

This study contributes to the empirical and experimental literature studying the reaction to incentives (for overviews see Charness and Kuhn, 2011; Lazear and Oyer, 2012; Camerer and Weber, 2013). Since, for example, the work of Nalbantian and Schotter (1997) and many following studies, this literature examines how incentive systems should be designed to induce high performance, without causing negative side effects. The overall finding is that (monetary) incentives change behavior, yet sometimes evoke possible dysfunctional responses (e.g., Asch, 1990; Ordóñez *et al.*, 2009; Gneezy *et al.*, 2011; Larkin, 2014). We demonstrate that not only the incentive side of the problem has to be taken into account, but that the opportunity cost side plays a crucial part which is often neglected.

Methodologically, our paper adds to the literature of real effort experiments, which are *"considered to be a better match to the field environment."* (Charness and Kuhn, 2011). These kind of experiments have been used to study such diverse phenomena as gender effects in competition (Niederle and Vesterlund, 2007), office politics (Carpenter *et al.*, 2010), and sorting into incentive schemes (Dohmen *et al.*, 2011). So far, most of the experimental literature using real effort experiments has considered fixed time environments.[2] By the nature of those experiments, performance changes can only be due to a change in the explicit costs of effort. Implicit costs play only a minor role in those experimental procedures, since subjects have to stay in the lab for a fixed time.

Our experiment is not the first attempt to integrate outside options for subjects (e.g., Mohnen *et al.*, 2008; Charness *et al.*, 2010; Corgnet *et al.*, 2015a,c; Koch and Nafziger, 2016), but only few studies manipulate them. The paper closest to ours is by Corgnet *et al.* (2015c). They study the effect of piece-rate and team incentives while varying the access to one real-leisure option, i.e. internet browsing. They find that effort under piece-rate incentives does not differ with and without a real-leisure alternative but under team based incentives. The present study differs from their paper in two aspects. First, we vary the implicit costs in various ways and demonstrate that this makes a difference, even for piece-rate incentives. Second, the focus of our paper is on individual incentives studying two piece-rate and two bonus schemes.

[2] One notable exception to this is Noussair and Stoop (2015), who use the time spend in the laboratory as a medium for reward. There, higher payoffs lead to shorter time in the lab.

Our results demonstrate that the effectiveness of incentive schemes depends crucially on the work environment. This helps to explain why in some work environments incentives might not change behavior. This non-responsiveness is unrelated to the incentive side of the problem, but simply due to the absence of implicit effort costs or to low opportunity costs.[3] If individuals face (nearly) no costs for their effort or behavior, the corresponding behavior might be performed independently of the incentive structure. Managers, who are able to control the opportunity costs of effort directly should take this into account and consider this part of the work environment more closely. Yet, even if the management is not able to control the costs of effort directly, it has to take into account that the behavioral responses to the incentive schemes depends on the given work environment. Thus, our results show the importance of taking implicit as well as explicit costs into account, when studying the behavioral response to incentive schemes or implementing them in practice.

The remainder of the paper is organized as follows. In section 2 we describe the design of our experiment and provide some behavioral hypotheses. Section 3 presents the results of the experiments. We conclude in section 4.

2 Design

Opportunity costs are the sum of implicit and explicit costs. In our computerized real-effort experiment we keep the explicit costs fixed while manipulating incentives and implicit costs between treatments. Based on the slider task by Gill and Prowse (2012) subjects had to adjust sliders ranging from 0 to 100 to the middle (50).[4] Each screen had 5 sliders and adjusting all 5 sliders yielded one point. The number of points a subject already acquired was displayed on the screen (see Figure 4 in the Appendix for a screenshot) and later used to calculate the payments. This task is constant in all treatments and the effort to move the slider represents the explicit cost part of the opportunity costs. The experiment consisted of 4 stages and implicit effort costs were manipulated in the third stage.

In the first stage, subjects work on the real effort task for 5 minutes without any monetary incentives. This stage serves two purposes: First, subjects learn the difficulty of the task and can form accurate expectations about the effort costs, and secondly, it provides an ability measure which is not influenced by the incentive scheme variation. In the second stage subjects received treatment specific instructions and were informed about

[3] In addition, our study helps to explain why incentives might not change behavior in real effort experiments utilizing the slider task (Araujo *et al.*, 2016).

[4]Subjects could only use the computer mouse. Keyboard and mouse-wheel were disabled.

Table 1: TREATMENTS AND DESCRIPTION OF THE 4×3-DESIGN

Treatment Name	Incentives Scheme	Work Enviroment
PIECERATE-LOW FIX	€0.02 per finished screen	No outside option, fixed duration of 40 minutes
PIECERATE-LOW INET	€0.02 per finished screen	Internet allowed, fixed duration of 40 minutes
PIECERATE-LOW FREE	€0.02 per finished screen	Free to leave, maximum duration of 40 minutes
PIECERATE-HIGH FIX	€0.1 per finished screen	No outside option, fixed duration of 40 minutes
PIECERATE-HIGH INET	€0.1 per finished screen	Internet allowed, fixed duration of 40 minutes
PIECERATE-HIGH FREE	€0.1 per finished screen	Free to leave, maximum duration of 40 minutes
BONUS-EASY FIX	€5 after 50 finished screens	No outside option, fixed duration of 40 minutes
BONUS-EASY INET	€5 after 50 finished screens	Internet allowed, fixed duration of 40 minutes
BONUS-EASY FREE	€5 after 50 finished screens	Free to leave, maximum duration of 40 minutes
BONUS-HARD FIX	€10 after 100 finished screens	No outside option, fixed duration of 40 minutes
BONUS-HARD INET	€10 after 100 finished screens	Internet allowed, fixed duration of 40 minutes
BONUS-HARD FREE	€10 after 100 finished screens	Free to leave, maximum duration of 40 minutes

the subsequently applied incentives. The dependent variable – output, i.e. number of completed screens – was obtained in the third stage of the experiment. In this stage subjects had to work on the real effort task for a maximum of 40 minutes. The exact implementation of this stage depended on the treatment. In the fourth stage subjects had to answer a short questionnaire, including socio demographics, the ten item version of the Big Five personality measure (Rammstedt and John, 2007), cognitive reflection test (Frederick, 2005), and general risk attitude (Dohmen *et al.*, 2011).

Treatments were implemented following a full 4×3 factorial design. Table 1 summarizes the implemented treatments. In the first treatment dimension we varied the incentives by implementing four different incentive schemes: two different piecerates (LOW or HIGH) and two different bonus schemes (EASY or HARD). In the piece-rate treatments subjects received a fixed payment for each successfully completed screen. In PIECERATE-LOW subjects received €0.02 per finished screen, in PIECERATE-HIGH €0.1. In the two bonus treatments subjects received a bonus conditional on reaching a pre-specified target.[5] In BONUS-EASY subjects received a €5 bonus if they reached the target of 50 screens. This is a relatively easy target, which most subjects could reach and in fact did. In BONUS-HARD subjects received a bonus of €10 if they reached the target of 100 screens. This target was set on purpose very high and no subject managed to reach the target. The size of the bonuses were chosen such that they equated the earnings of a subject in the high piece-rate treatment with the same number of completed screens. Thus, a subject with 50 completed screens would earn the same in BONUS-EASY and PIECERATE-HIGH

[5]The targets were chosen based on pilots run with the real-effort task

and a subject with 100 completed screens would earn the same in Bonus-Hard and Piecerate-High.

In the second treatment dimension, we manipulated the implicit costs by implementing three different work environments. First, in Fix, we implemented a fix time procedure, in which subjects had to stay for 40 minutes at the computer without any offered leisure alternatives.[6] We manipulated the implicit costs by implementing two environments with alternative activities for the subjects. In the Inet-environment, subjects had to stay in front of their computer, but were allowed to use the Internet Explorer during the working phase of the experiment. Subjects had to remain in the laboratory for the full time, but could surf the web instead of working on the task. This was implemented with a button on the real-effort screen, which would open the Internet Explorer and hide the real-effort task. Subjects could not work on the real effort task and use the Internet Explorer at the same time. However, subjects could always close the Internet Explorer and press a button to return to the real effort task. This allows us to record how much time subjects spent on the real effort task or in the internet. In the treatment condition Free, subjects could adjust their working time between 0 and 40 minutes by stopping to work on the real effort task whenever they wanted. The screen in the working stage included a leave button. Pressing the button lead to the questionnaire and subjects could leave the cubicle to get their payments after finishing it. Payments were made based on the number of screens at the time the subject stopped working.

The experiments were conducted at the BonnEconLab of the University of Bonn and subjects were recruited via hroot (Bock *et al.*, 2014). Upon arrival, subjects were seated in cubicles with curtains and blinds up to the ceiling which prevented the subjects from observing anything outside their cubicle. We conducted eight regular sessions with the Fix- and Inet-treatments and slightly adjusted the implementation in the Free-treatments. In the Free-treatments, subjects were invited to the lab on given days at any time between 10am and 4pm. This procedure ensured that subjects would not know the duration other subjects spend working on the task preventing possible spillovers. Furthermore, this procedure removed fixed beliefs about the duration of the experiment. In all treatments subjects received their payments individually in a separate room. The experiment itself was computerized using Ztree (Fischbacher, 2007).

For each treatment we gathered approx. 24 observations. In total 288 Subjects participated, with 56% of subjects being female and an average age of 23.5 years.[7] A session lasted on average 75 minutes for Fix and Inet and individual sessions in Free lasted

[6] The use of mobile phones was forbidden in all treatments.

[7] Neither gender nor age differ significantly between the three different work environments ($p = 0.491$ and $p = 0.63$, both Kruskal-Wallis test). In addition, ability, as measured in the first stage, does not differ significantly between the three work environments (p = 0.773, Kruskal-Wallis test).

between 20 and 75 minutes. All subjects received a show-up fee of €10 and additional earnings from the real effort task. Subjects earned on average €12.65, ranging from €10 in the BONUS-HARD-treatments to an average of €15.99 in the treatment PIECERATE-HIGH FIX.

2.1 Behavioral Hypotheses

In a simple theoretical framework the effort level would be chosen by maximizing

$$u(e) = \bar{w} + b(y) + I\delta(y) - c(e, i),$$

with a production technology $y = f(e)$, a fixed wage $\bar{w}$ (in our experiment the show-up fee), a performance dependent payment $b(y)$ (either piecerate or bonus), the intrinsic motivation $I\delta$ [8] , and some costs depending on the explicit costs of effort e and the implicit costs i.[9] In what follows we present our behavioral predictions focusing on the intuition, but we discuss the underlying framework in more detail in the appendix.

With our work environment manipulation we increase the marginal effort costs in INET and FREE compared to FIX.[10] Thus, as the marginal costs of effort increases, output should be higher in FIX than in the other two work environments.

Hypothesis 1: *We expect higher output in* FIX *than in* INET *and* FREE.

Let us now consider the differences between piece-rate treatments. Subjects provide effort as long as the marginal benefits from the piece-rate payment and the intrinsic motivation to perform the task are higher as the marginal costs of effort. This point is reached sooner in PIECERATE-LOW than in PIECERATE-HIGH, due to the lower marginal benefits in PIECERATE-LOW, leading to higher outputs in the latter one. This holds for the comparison of all piece-rate treatments, within a work environment.

Hypothesis 2: *We expect higher output in* PIECERATE-HIGH *than in* PIECERATE-LOW.

[8] Following the approaches by Murdock (2002) and James (2005) δ represents the agent's intrinsic motivation for the work and I is an indicator function which is $I = 1$ if the agent is intrinsically motivated and $I = 0$ if not.

[9] This includes for example versions of $c(e, i)$ like in Koch and Nafziger (2016), where $i \times \tilde{c}(e)$ where the parameter i differs between work environments, i.e., effort costs increase when alternative actions are present. Therefore $i_{Fix} \leq i_{Surf} \leq i_{Leave}$. This could also incorporate a version of $c(e, i)$, where implicit costs are modeled as utility of leisure, but leisure is negatively related with effort, i.e. time, as in (Corgnet *et al.*, 2015c)

[10] Strictly speaking we do not have enough information to know that implicit costs are higher in FREE than in INET, but it seems reasonable to assume that subjects have more outside opportunities than just surfing the web in FREE.

For Bonus-Easy we would expect only few outputs above 50 as $\frac{\partial b(y)}{\partial y} = 0$ for any output above 50 and additional output would only be driven by workers for which the marginal intrinsic motivation would still be higher than the marginal costs. In Bonus-Hard we would expect very low output in general, since subjects should realize very early on in the experiment that they will not reach the target of 100 and thus marginal (monetary) benefits equal zero for all feasible outputs. Thus, output would again only be driven by workers for which the marginal intrinsic motivation is higher than the marginal costs.

We are agnostic about the differences between Bonus-Easy and the two piece-rate treatments. Predictions about performance differences would involve assumptions not only about the exact form of the cost of effort function, but also the intrinsic motivation. However, since reaching the target of 100 screens in Bonus-Hard is not feasible we can predict that in both piece-rate treatments output should be higher than in Bonus-Hard. This is due to the fact that monetary incentives are basically not present in the Bonus-Hard and therefore incentives are higher in the two piece-rates treatments.

Hypothesis 3: *We expect higher output in* Bonus-Easy *than in* Bonus-Hard. *The output in* Bonus-Easy *should be centered around 50 screens. Furthermore, we expect higher output in both piece-rate treatments than in* Bonus-Hard

3 Results

In the following we demonstrate that implicit costs have a significant impact on work output. We will first discuss the influence of those within an incentive scheme before demonstrating that implicit costs influence the comparisons between incentive contracts. Thereafter, we will take a closer look at the usage of the offered outside option and the influence of non-cognitive traits on behavior.

The impact of implicit costs on output

Table 2 gives the summary statistics for the outputs in all treatments. In general, we observe significantly higher output in the FIX-treatments with no offered outside option compared to the treatments with offered outside options. On average the output in the FIX-treatments is 22.4% higher than in the INET-treatments ($p < 0.001$, two-sided Mann-Whitney u-test) and 34% higher than in the FREE-treatments ($p < 0.001$). In addition the average output in the INET-treatments is 9.6% higher than in the FREE-treatments but this difference turns out to be insignificant ($p = 0.298$). Already based on this general look at the data we can conclude that implicit costs in general influence the output negatively, which is in line with our predictions.

Result 1: *On average, subjects produce significantly higher outputs in* FIX *than in* INET *and* FREE. *The offered outside options – resulting in increased implicit costs – influence the output.*

Table 2: SUMMARY STATISTICS OF OUTPUTS

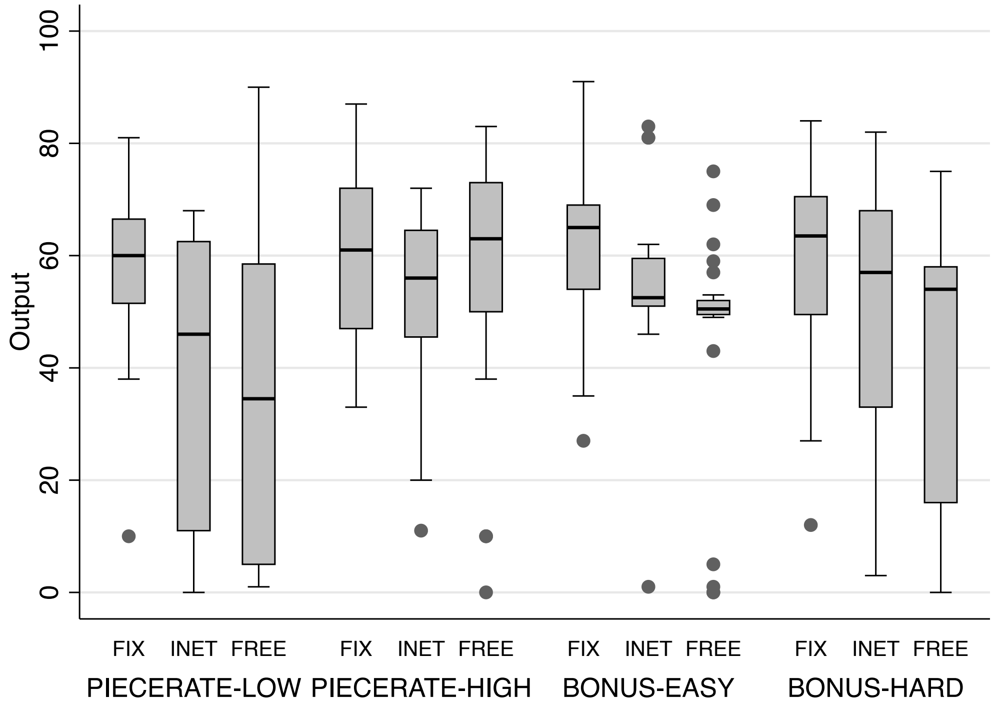
		OVER ALL INCENTIVES	BY INCENTIVE SCHEME PIECERATE-LOW	PIECERATE-HIGH	BONUS-EASY	BONUS-HARD
FIX	Mean	59.76	58.54	59.87	61.86	59.04
	SD	15.14	14.24	13.84	14.29	18.36
	N	92	24	23	21	24
INET	Mean	48.8	38.25	53.00	55.00	48.96
	SD	20.91	25.19	15.01	15.37	23.22
	N	95	24	24	24	23
FREE	Mean	44.53	35.46	56.96	44.96	40.40
	SD	25.21	28.02	22.56	21.00	24.98
	N	98	24	25	24	25

SD: standard deviation, N: number of independent observations

In the following we turn to differences within each incentive scheme. Figure 1 shows boxplots of the output for each incentive scheme and each implicit cost settings.

In PIECERATE-LOW we observe a decline of effort between FIX and both INET and FREE. Output decreases on average by 39.4 % in FREE and by 34.7 % in INET when sub-

Figure 1: Boxplot of Outputs in the different Treatments



Bold lines give the median outputs, boxes the 25th and 75th quartiles, and whiskers the 1.5xIQR.

jects face increased implicit effort costs, compared to Fix. Pairwise comparison between Piecerate-Low Fix and the other two treatments reveals this difference to be statistically significantly for both pairwise comparisons (Inet: $p = 0.014$; Free: $p = 0.003$, both two-sided Mann-Whitney u-test).[11] Increasing the implicit costs leads to lower, but similar outputs in both Piecerate-Low Inet and Piecerate-Low Free ($p = 0.65$)[12]. We can conclude that in both environments the marginal costs of effort increases at least for some subjects due to increased implicit costs. Effort for those subjects decreases and, therefore, the average output decreases.[13]

However, this does not hold in general for all monotone incentive schemes. The performance of subjects in the Piecerate-High treatments is similar in all implicit costs treatments – Free, Inet and Fix. All pairwise comparisons reject the existence of statistically significant differences (all $p > 0.155$). When subjects face high incentives, high effort is induced, resulting in high output in all of our work environments. For high in-

[11] The results presented in this section can be confirmed with OLS regressions controlling for ability, gender, and age. Test results are in line with the presented non-parametric results. All significant results remain on conventional significance levels with the additional controls, see Table 4 in the Appendix.

[12] If not reported otherwise, all p-values in the paper use the two-sided Mann-Whitney u-test.

[13] These significant differences in output between the work environments are mostly driven by the usage of the offered outside option. Refer to Appendix A.3 for some analyses of the time spend working on the task.

centives, the increased implicit costs in our experiments seem not to be high enough to induce statistically significant lower effort. Nevertheless, observed performances decrease slightly and are generally in line with our hypothesis. The small differences we observe could be for example due to the fact that high marginal incentives induce an effort level, which is on the steep part of the marginal effort curve for all environments, resulting in only small differences between treatments.

Result 2: *Subjects in* PIECERATE-HIGH *facing high monotone incentives do not respond significantly to the changes in implicit costs. In contrast, subjects in* PIECERATE-LOW *facing low monotone incentives respond to changes in implicit costs resulting in significantly lower outputs in* INET *and* FREE *compared to* FIX.

Similar to the PIECERATE-LOW treatments and in line with our predictions, we observe a decline of output for both bonus-based incentive schemes with increased implicit costs. For BONUS-EASY the output in FIX 12.5% is higher than in INET ($p = 0.049$) and 37.6% higher than in FREE ($p = 0.0019$). Additionally, we observe a 22.3% higher output in BONUS-EASY INET compared to BONUS-EASY FREE, which again is significantly different ($p = 0.0171$). In BONUS-EASY FREE the performance distribution collapses around 50 screens. This is due to the majority of subjects who stopped once they reached the threshold for the bonus.[14] Only a few subjects worked more than necessary and some subjects stopped early, performing poorly. Interestingly, this sharp decline in effort provision beyond 50 cannot be observed in BONUS-EASY INET. However, those differences in outputs do not translate into significant differences in the number of subjects, who earned the bonus. In FIX 81% of subjects reached the target of 50 screens, in INET 92% and in FREE 75% (all pairwise comparisons p>0.245, two-sided Fisher's exact test).

A declining output pattern can also be observed in BONUS-HARD. In FREE, the output distribution is shifted downwards, compared to FIX, and has a longer lower tail (again compare Figure 1). The lower average output is driven by subjects who stop working before the time is up. This difference in performance is also statistically significant ($p < 0.01$, two-sided Mann-Whitney u-test). The output in INET is between FIX and FREE and does not differ statistically from both (for both $p > 0.215$). No subject in the BONUS-HARD-treatments is able to reach the target of 100 screens; the highest output was 84.

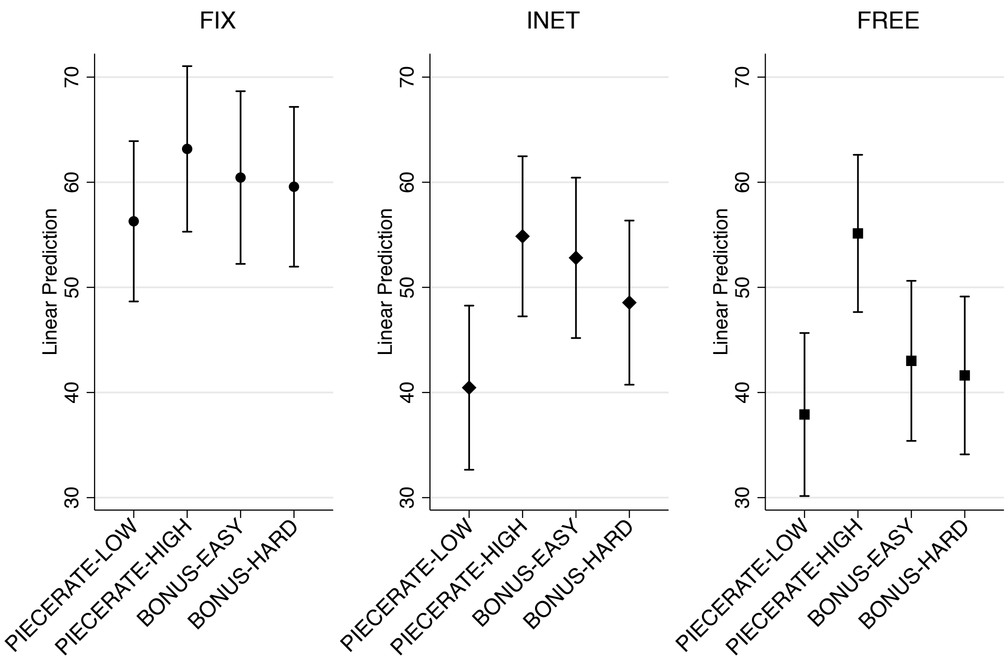
Result 3: *Subjects in the two bonus treatments respond to changes in implicit costs. In* BONUS-EASY *the two treatments with increased implicit costs,* INET *and* FREE, *result in significantly lower output than* FIX. *In* BONUS-HARD *subjects produce significantly lower outputs in* FREE *compared to* FIX.

[14] Again, the different outputs result from different durations spend working on the task. Refer to Appendix A.3 for additional analyses.

Implicit costs and the comparison between incentive schemes

So far we have demonstrated that implicit costs can influence the output even if incentives for the (marginal) output are fixed. In a next step we will demonstrate that implicit costs influence the comparison of incentive schemes. Again, refer to Table 2 for descriptive statistics. In addition, Figure 2 presents the estimated output in all treatments after controlling for ability, gender and age.[15]

Figure 2: PREDICTED OUTPUT WITH 95% CIS



Estimates are based on linear regression controlling for subjects' ability, gender, and age. For results see table 4 in appendix.

We start by analyzing the output of subjects in the real effort task for the FIX-treatments. In all four incentive schemes, PIECERATE-LOW, PIECERATE-HIGH, BONUS-EASY and BONUS-HARD, subjects average output is close to 60 points, with considerable variation between subjects. All pairwise comparisons of average outputs fail to reject the null hypothesis of equal means between any two incentives schemes in FIX (all comparisons $p > 0.4$, two-sided Mann-Whitney u-test). As Figure 2 reveals, this non-result remains robust to additional controls for ability, gender, and age (all $p \geq 0.22$, Wald test). Thus, subjects' effort in FIX is not differently influenced by the diverse incentive schemes.

Comparing output across incentive schemes for INET, we observe some variations across incentive schemes. The larger incentives in PIECERATE-HIGH result in a significantly

[15]Refer to Table 4 in the Appendix for the estimation results the figure is based on.

higher output of 38.6 % compared to PIECERATE-LOW ($p = 0.085$, Mann-Whitney u-test). While average outputs in BONUS-EASY and BONUS-HARD are 43.8% and 28% larger than in PIECERATE-LOW they do not differ significantly ($p > 0.11$, Mann-Whitney u-test). However, adding controls does not only turn the difference between PIECERATE-HIGH and PIECERATE-LOW into highly significant ($p < 0.01$, Wald test) but also turns the difference between PIECERATE-LOW and BONUS-EASY to be significant ($p = 0.027$, Wald test). All other comparisons remain insignificant ($p > 0.154$, Wald test).

If we compare the incentive schemes within the working environment FREE we are able to observe even stronger differences in the response to incentives. The output in PIECERATE-HIGH is 60.6% higher than in PIECERATE-LOW, 26.7% higher than in BONUS-EASY, and 40.1% higher than in BONUS-HARD. All three comparisons are statistically significant when applying non-parametric tests ($p < 0.01, p = 0.014$, and $p = 0.026$, two-sided Mann-Whitney u-test), as well as, using parametric tests controlling for ability, gender and age ($p < 0.01, p = 0.026$, and $p = 0.013$, Wald test). The average outputs in the bonus treatments are higher than in PIECERATE-LOW, yet, all other pairwise comparisons remain insignificant when controlling for ability, gender and age (all $p > 0.356$, Wald test).

The comparison of the inventive schemes in FIX is in sharp contrast to our predictions. In the absence of implicit costs incentives fail to induce significantly different responses. In INET and FREE implicit effort costs increase and the subjects' responses to the incentive schemes are generally in line with our predictions. To summarize:

Result 4: *The comparison of incentive schemes depends on the implicit cost. While no significant output differences are observed in* FIX*, significantly different outputs between the incentive schemes are observed in* INET *and* FREE*.*

Elasticity of Output

A different way to investigate the reaction to changed incentives is to calculate the elasticity of the output in all three work environments with regard to the piecerates. In PIECERATE-HIGH marginal incentives are higher for each additionally produced screen than in PIECERATE-LOW. Thus, from a purely incentive theory perspective we would on average expect higher outputs in PIECERATE-HIGH compared to PIECERATE-LOW, i.e., a positive elasticity. Table 3 gives the resulting elasticities when regressing the logarithm of the piecerate on the logarithm of the output. For FIX we observe an elasticity close to zero.[16] This finding is in line with our previous null result of subjects' output not responding to changes of the piecerate. For both, INET and FREE, we observe significantly

[16] It is worth noting that our estimated output elasticity of 0.0298 in FIX is very similar to the 0.025 estimated by Araujo *et al.* (2016).

larger and positive elasticities compared to Fix ($p = 0.037$ and $p < 0.01$, Wald test). In Inet increasing the piecerate by 1% would increase the outputs by 0.35% and in Free by 0.65%. These two elasticities do not differ significantly from each other ($p = 0.208$, Wald Test).

Table 3: Elasticities

	Fix	Inet	Free
ln(Piecerate)	0.0298	0.3498**	0.6511***
	(0.0605)	(0.1419)	(0.1962)
Constant	4.1336***	4.7123***	5.4811***
	(0.1951)	(0.4532)	(0.6299)
N	47	47	48
R^2	.0054	.12	.19

* $p < .1$, ** $p < 0.05$, *** $p < 0.01$
standard errors in parentheses

These results show that implicit costs induced by different work environments matter. While subjects' output does not respond to increased incentives in Stay, we are able to observe reactions in outputs to increased incentives in both working environments with higher implicit costs. While the response in Inet and Free is positive it is still inelastic (i.e., below 1%).[17]

Result 5: *In* Fix *the elasticity of the output is not distinguishable from zero. In both* Inet *and* Free *we observe a positive yet inelastic response of the output to increased incentives.*

Supplementary Analyses

In the following we take a closer look at individuals characteristics and personality traits and their impact in the different work environments. We will first start with subjects' ability as measured by the number of finished screens in the first stage.[18]

In all three implicit costs settings the output and ability measure are significantly correlated, however, the strength of these correlations differs substantially (Fix: $\rho = 0.6982$ with $p < 0.001$; Inet: $\rho = 0.4379$ with $p < 0.001$, Free: $\rho = 0.2807$ with $p = 0.005$, Spearman rank correlation). This translates into different output predictions based on the ability.

[17] Although this is an inelastic response, it would be surprising to find a higher output elasticity. Given that subjects can only solve a limited amount of sliders within 40 minutes and therefore elastic output responses are nearly impossible to observe given the baseline output in Piecerate-Low.

[18] Ability does not differ significantly between Fix, Inet, and Free ($p = 0.773$, Kruskal-Wallis test).

However, as ability does not differ between treatments, it can not be ability itself, which can explain the differences between treatments. In fact, these differences in output result from a substantial fraction of subjects using the offered outside options when available: 37% in Inet and 46% in Free. Thus, subjects in both work environments spend significantly less than 40 minutes working on the task (both $p < 0.001$ sign-rank test, two sided). In Inet subjects work on average 34 minutes on the task and in Free 29 minutes ($p = 0.044$, Mann-Whitney u-test).[19]

In a last step we use our questionnaire data to investigate the characteristics of those subjects that used the offered outside option. Along with gender and age, we also elicited general risk attitudes, personality traits, and cognitive ability. We elicited the general risk attitudes (Dohmen *et al.*, 2011) as the two Bonus-treatments involve the risk of investing effort without reaching the target. To elicit personality traits we administered the 10-item version of the Big 5 (Rammstedt and John, 2007). Of the personality traits we are particularly interested in the effect of conscientiousness, which has been linked to increased job performance (e.g., Barrick and Mount, 1991).[20] In fact, research on non-cognitive skills suggests that conscientiousness predicts educational attainment and labor market outcomes as strongly as cognitive ability (Heckman and Kautz, 2012). As an additional measure, we implemented the CRT, a cognitive reflection test by Frederick (2005). A recent paper by Corgnet *et al.* (2015b) shows for a setting similar to our Inet-environment that higher cognitive reflection reduces leisure activities.

Figure 3 shows the estimated impact of these traits and characteristics on the time worked on the task while controlling for incentive scheme, subjects' ability, gender, and age.[21] Neither in the Inet-treatments nor the Free-treatments general risk attitudes are significantly correlated with the time worked on the task (both $0 > \rho > -0.1$ and $p > 0.32$, Spearman rank correlation). As expected, conscientiousness, the desire to do a task well, is positively correlated with the time subjects work on the task in the Inet- ($\rho = 0.1962, p = 0.057$, Spearman rank correlation) and the Free-treatments ($\rho = 0.2052, p = 0.043$, Spearman rank correlation).

In Inet, cognitive reflection is not significantly correlated with the time spend working in the task ($\rho = -0.0473, p = 0.6493$ Spearman correlation). However, in Free higher cognitive reflection is significantly correlated with less time worked on the task ($\rho = -0.2619, p = 0.0092$).[22] Taking a closer look reveals that the negative correlation is

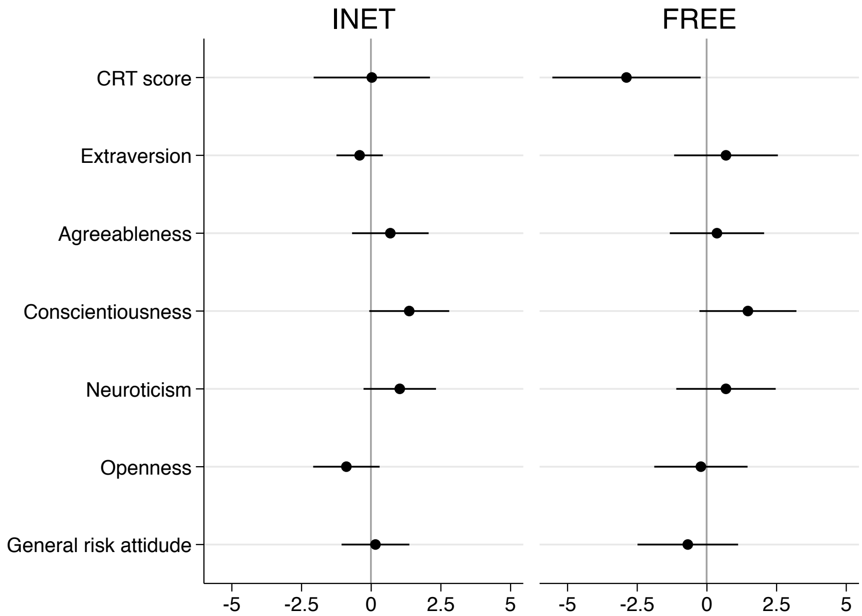
[19]See appendix A.3 for some more analyzes on the differences in time spend working on the task

[20]The American Psychology Association defines conscientiousness as "the tendency to be organized, responsible, and hardworking".

[21]For coefficients of the OLS regressions see Table 5 in the appendix.

[22]This result is in contrast to the findings by Corgnet *et al.* (2015b). However, the summation task implemented in Corgnet *et al.* (2015b) might be cognitively more demanding than the slider task. In this case, cognitive ability would play a more important role for effort costs than in our experiment.

Figure 3: Coefficients for time spend on task with 95% CIs



Estimates are based on linear regression controlling for incentive schemes, and subjects' ability, gender, and age. For coefficients see Table 5 in the appendix.

stronger in the Bonus-treatments ($\rho = -0.3228, p = 0.0237$) compared to the Piecerate-treatments ($\rho = -0.1901, p = 0.1908$). This suggests, that subjects with higher cognitive reflection realize (sooner) that there are no monetary incentives to work beyond 50 screens in Bonus-Easy and to work at all in Bonus-Hard with an unrealistic target. Figure 3 demonstrates that the described effects remain robust when controlling for them simultaneously while also controlling for the incentive schemes, subjects' ability, gender, and age.[23]

Result 6: *Conscientiousness is significantly correlated with more time spend working on the task, while increased cognitive reflection leads to less time spend on the task if subjects have the opportunity to leave the lab.*

[23] Table 5 in the appendix shows that Conscientiousness and CRT predict not only the time spend on the task, but also if a subject decided to leave the lab in Free.

4 Conclusion

In this paper we investigate how work environments with different implicit costs influence the effectiveness of non-monotone and monotone incentive schemes. We exogenously vary the implicit effort costs between work environments by offering real-leisure alternatives and compare the performance of subjects in two piece-rates and two bonus schemes. We observe that incentive contracts and opportunity costs interact in a non-trivial manner. Performance reacts significantly to changes in implicit effort costs under low-powered piece-rate and target-based bonus contracts, but not under a high piece rate contract. In addition, comparisons between the incentive schemes depend crucially on the work environment, i.e., the implicit costs. In our FIX environment subjects do not react to the incentive contracts. However, in our environments with increased implicit costs (INET and FREE) subjects' performances differ between incentive schemes and positive elasticities for the output are observed.

With high monotone incentives, as in our PIECERATE-HIGH treatments, workers' outputs are relatively stable regardless of the work environment and its implicit costs. However, with low monotone incentives as in our PIECERATE-LOW treatments, workers' output are negatively impacted by the implicit costs. More generally, we observe a positive output elasticity only in those work environments with increased implicit effort costs. With respect to non-monotone incentive schemes, our results suggest that the effect of bonus schemes depends on the opportunities of workers to allocate their time, i.e. which effort costs they face. Our results in BONUS-EASY suggest that targets in work environments with high implicit effort costs induce behavior such that targets are closely matched, but not exceeded. In addition, targets like in BONUS-HARD, might increase the number of workers who drop out of the task once they realize that the target is unrealistic. However, in the environment with low implicit costs, we observe effort beyond incentivized points. This might be unexpected, but is similar to the fixed bonus treatments reported for example in (DellaVigna and Pope, 2016), where experts also fail to forecast this effort provision beyond the incentivized point.

Our results have implications for the use and design of incentive schemes within organizations. The management has many means to affect workers' behavior and every aspect of an organization can be used as a parameter to obtain desired outcomes (Roberts, 2007). Organizations should recognize that implicit costs are a relevant parameter in addition to monetary incentives which they might want to adjust. Even if the work environment is fixed and cannot easily be influenced by the management it is important to take these implicit costs into account when implementing traditional incentive schemes. Our results are also interesting in light of the recent discussion of more workplace flexibility and home-offices. Those changes at the work place might not change the incentive structure and might not even change the supervision or control of individuals. However, those changes are likely to change the implicit effort cost structure. The results of this paper suggest that incentives used in traditional work environments might perform differently in those newer environments.

References

Araujo, F. A., Carbone, E., Conell-Price, L., W. Dunietz, M., Jaroszewicz, A., Landsman, R., Lamé, D., Vesterlund, L., Wang, S. W. and Wilson, A. J. (2016). The slider task: an example of restricted inference on incentive effects. *Experimental Economics*, **2** (1), 1–12.

Asch, B. (1990). Do incentives matter? The case of navy recruiters. *Industrial & Labor Relations Review*.

Barrick, M. R. and Mount, M. (1991). The big 5 personality dimensions and job-performance: a meta-analysis. *Personnel psychology*, **44** (1), 1–26.

Bock, O., Baetge, I. and Nicklisch, A. (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review*, **71**, 117–120.

Bradler, C., Dur, R., Neckermann, S. and Non, A. (2016). Employee recognition and performance: A field experiment. *Management Science*, **3**.

Camerer, C. F. and Weber, R. A. (2013). Experimental organizational economics. *The Handbook of Organizational Economics*, pp. 213–262.

Carpenter, J. (2016). The labor supply of fixed-wage workers: Estimates from a real effort experiment. *European Economic Review*, **89**, 85–95.

—, Matthews, P. H. and Schirm, J. (2010). Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*, **100** (1), 504–417.

Charness, G. and Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab? *Handbook of labor economics*, **4**, 229–330.

—, Masclet, D. and Villeval, M. C. (2010). Competitive preferences and status as an incentive: Experimental evidence. *Groupe d'Analyse et de Théorie Économique Working Paper*, (1016).

Cohn, A., Fehr, E. and Goette, L. (2015). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, **61** (8), 1777–1794.

Corgnet, B., Gómez-Miñambres, J. and Hernán-González, R. (2015a). Goal Setting and Monetary Incentives: When Large Stakes Are Not Enough. *Management Science*, **61** (12), 2926–2944.

—, Hernán-González, R. and Mateo, R. (2015b). Cognitive Reflection and the Diligent Worker: An Experimental Study of Millennials. *PLoS ONE*, **10** (11).

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

Deci, E. L. (1971). The effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, **18**, 105–115.

DellaVigna, S. and Pope, D. (2016). What Motivates Effort? Evidence and Expert Forecasts. *NBER Working Paper*, **22193**.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, **9** (3), 522–550.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, **96** (5), 1611–1630.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, **10** (2), 171–178.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*.

Gill, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *The American economic review*, pp. 469–503.

Gneezy, U., Meier, S. and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, **25** (4), 191–210.

Goerg, S. and Kube, S. (2012). Goals (th)at Work, Goals , Monetary Incentives , and Workers ' Performance.

Goerg, S. J., Kube, S. and Zultan, R. (2010). Treating equals unequally: Incentives in teams, workers' motivation, and production technology. *Journal of Labor Economics*, **28** (4), 747–772.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, **19**, 451–464.

Herweg, F., Müller, D. and Weinschenk, P. (2010). Binary payment schemes: Moral hazard and loss aversion. *American Economic Review*, **100** (5), 2451–2477.

Holmstrom, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pp. 303–328.

— and — (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, **7**, 24–52.

James, H. S. (2005). Why did you do that? an economic examination of the effect of extrinsic compensation on intrinsic motivation and performance. *Journal of Economic Psychology*, **26** (4), 549–566.

Koch, A. K. and Nafziger, J. (2011). Self-regulation through goal setting. *Scandinavian Journal of Economics*, **113** (1), 212–227.

— and — (2016). Gift exchange, control, and cyberloafing: A real-effort experiment. *Journal of Economic Behavior & Organization*, **131, Part A**, 409–426.

Kosfeld, M. and Neckermann, S. (2011). Getting more work for nothing? symbolic awards and worker performance. *American Economic Journal: Microeconomics*, **3**, 86–99.

Larkin, I. (2014). The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales. *Journal of Labor Economics*, **32** (2), 199–227.

Lazear, E. P. and Oyer, P. (2012). *Personnel Economics*, Princeton University Press, pp. 479–519.

Lazear, P. E. (2000). Performance Pay and Productivity. *American Economic Review*, **90** (5), 1346–1361.

Mohnen, A., Pokorny, K. and Sliwka, D. (2008). Transparency, inequity aversion, and the dynamics of peer pressure in teams: Theory and evidence. *Journal of Labor Economics*, **26** (4), 693–720.

Murdock, K. (2002). Intrinsic motivation and optimal incentive contracts. *RAND Journal of Economics*, **33** (4), 650–671.

Nalbantian, H. R. and Schotter, A. (1997). Productivity under group incentives: An experimental study. *The American Economic Review*, pp. 314–341.

Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, **122** (3), 1067–1101.

Noussair, C. and Stoop, J. (2015). Time as a medium of reward in three social preference experiments. *Experimental Economics*, **18** (3), 442–456.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D. and Bazerman, M. H. (2009). Goals Gone Wild: The Systematic Side Effects of Overprescribing Goal Setting. *Academy of Management Perspectives*, **23** (1), 6–16.

Rammstedt, B. and John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German.

Roberts, J. (2007). *The modern firm: Organizational design for performance and growth*. Oxford university press.

Winter, E. (2004). Incentives and Discrimination. *American Economic Review*, **94** (3), 764–773.

A Appendix

A.1 Screenshot

Figure 4: Screenshot of Real Effort Screen

A.2 Additional Figures and Tables

Table 4: OLS Regression: Output in the Experiment

	(1)	(2)
Piecerate-Low × Inet	-16.56***	-15.83***
	(5.490)	(5.535)
Piecerate-Low × Free	-19.55***	-18.38***
	(5.487)	(5.532)
Piecerate-High × Fix	5.775	6.884
	(5.561)	(5.600)
Piecerate-High × Inet	-1.419	-1.428
	(5.497)	(5.494)
Piecerate-High × Free	-0.773	-1.155
	(5.408)	(5.412)
Bonus-Easy × Fix	4.690	4.160
	(5.657)	(5.695)
Bonus-Easy × Inet	-3.149	-3.478
	(5.462)	(5.456)
Bonus-Easy × Free	-13.58**	-13.28**
	(5.461)	(5.457)
Bonus-Hard × Fix	2.856	3.284
	(5.473)	(5.469)
Bonus-Hard × Inet	-7.187	-7.737
	(5.532)	(5.545)
Bonus-Hard × Free	-15.45***	-14.67***
	(5.422)	(5.463)
Ability	4.711***	4.940***
	(0.709)	(0.743)
Female		3.963
		(2.438)
Age		0.0261
		(0.233)
Constant	28.90***	24.57***
	(5.899)	(8.933)
N	285	284
R Squared	.28	.28

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Time spend working and probability using outside option

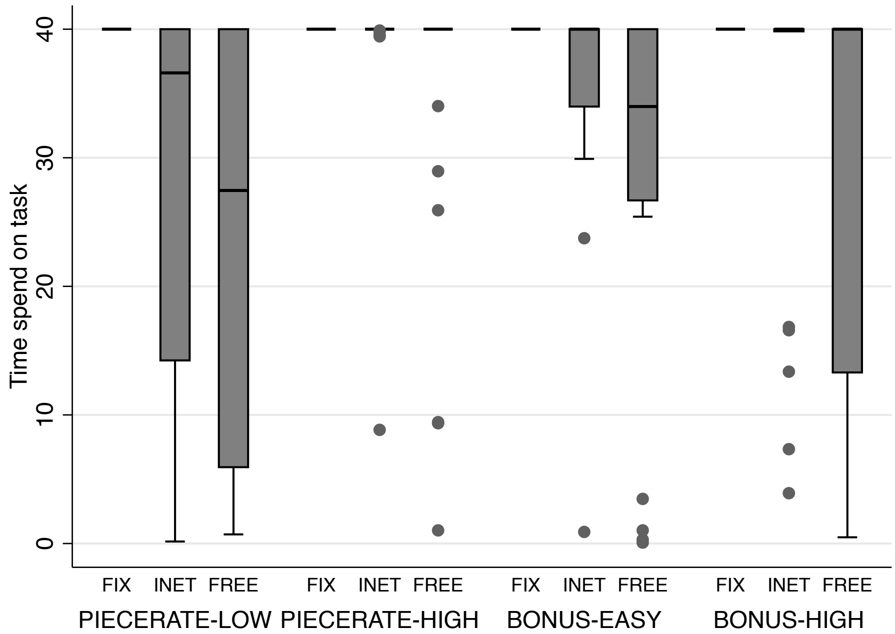
	Inet		Free	
	Time (OLS)	Use Inet (Logit)	Time (OLS)	Leave Lab (Logit)
CRT score	0.0216	-0.0113	-2.885**	0.455*
	(1.048)	(0.207)	(1.338)	(0.252)
Extraversion	-0.415	0.176	0.685	-0.201
	(0.418)	(0.131)	(0.936)	(0.154)
Agreeableness	0.687	-0.143	0.361	-0.0928
	(0.691)	(0.178)	(0.851)	(0.168)
Conscientiousness	1.365*	-0.152	1.472*	-0.415**
	(0.721)	(0.135)	(0.875)	(0.165)
Neuroticism	1.025	-0.0166	0.684	-0.249*
	(0.653)	(0.129)	(0.897)	(0.147)
Openness	-0.890	-0.108	-0.215	-0.0672
	(0.599)	(0.143)	(0.842)	(0.133)
General Risk	0.153	0.0626	-0.685	0.000616
	(0.611)	(0.124)	(0.907)	(0.158)
Constant	25.43*	1.010	11.34	6.475**
	(13.12)	(2.397)	(13.36)	(3.015)
Controls:	Yes	Yes	Yes	Yes
N	95	95	97	97
R Squared	.27	0.15	.2	0.25

Controls: ability, incentive scheme, age, female

A.3 Time used to work on the task

Our treatment variation gave subjects the possibility to adjust their effort at the extensive margin in Inet and Free by using a provided outside option. In the following we show that subjects actually used the outside option. In addition, we show that the time worked on the task (extensive margin) rather than the speed (intensive margin) results in the reported treatment differences.

Figure 5: Boxplot of time spend working on the task

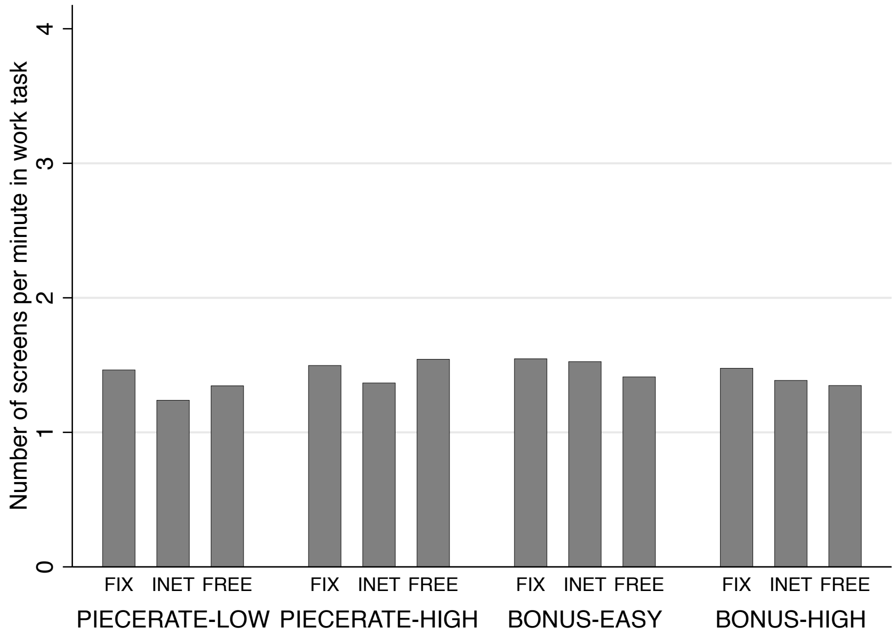


Over all treatments is the time spend working on the task significantly correlated with the output (r=0.6279, r<0.001). Figure 5 shows the distribution of time spend working on the task. Obviously, subjects worked on average significantly less than 40 minutes on the task in the Inet- and Free-treatments (both $p < 0.001$, Wilcoxon signed-rank test).

In a next step we adjust the output by the total time an individual worked on the task. If the time spend working on the task is the main driver of the treatment differences, we should not observe significant differences between treatments anymore. Figure 6 gives the result of this exercise. The previously reported significant differences between Fix, Inet, and Free turn insignificant when controlling for the time spend working on the task: In Piecerate-Low the p-value changes from $p = 0.006$ to $p = 0.278$, in Bonus-Easy from $p = 0.002$ to $p = 0.675$ and in Bonus-Hard from $p = 0.025$ to $p = 0.467$ (all Kruskal-Wallis test). Only in Piecerate-High turns the difference from insignificant ($p = 0.295$) to borderline significant ($p = 0.095$).

We additionally analyze the time subjects need to complete one screen to measure effort adjustments at the intensive margin. Table 6 gives the average and median time needed per screen. Testing the median time per screen reveals no signifiant differences in the medians (all $p > 0.12$, median test). Comparing the means gives no significant differences

Figure 6: Output, adjusted by working time



for Piecerate-Low, Bonus-Easy and Bonus-Hard (all $p > 0.365$, Kruskal-Wallis test). In Piecerate-High we observe a significant difference ($p < 0.1$).

Table 6: Time per Screen

		Piecerate-Low	Piecerate-High	Bonus-Easy	Bonus-Hard
Fix	average	49.00	42.41	41.66	49.47
	median	40.00	39.34	36.92	37.80
	standard deviation	41.31	10.93	13.98	35.50
	Observations	24	23	21	24
Free	average	52.87	39.14	40.36	47.17
	median	42.26	36.37	39.69	41.74
	standard deviation	32.23	9.44	8.15	18.51
	Observations	24	24	22	24
Inet	average	55.10	47.21	40.59	47.25
	median	42.26	36.37	39.69	41.74
	standard deviation	29.03	17.26	7.01	15.61
	Observations	23	24	24	23

A.4 Conceptual Framework

In a simple theoretical framework the effort level would be chosen by solving the following maximization problem.

$$\max_{e \geq 0} u(e) = \bar{w} + b(y) + I\delta(y) - c(e,i),$$

The production technology $y = f(e)$ translates effort to output, which we assume to be a continuously differentiable function with $f' > 0$ and $f'' < 0$. The fixed wage, i.e. a lump-sum payment, is represented by $\bar{w}$. The intrinsic motivation is represented by $I\delta(y)$, which indicates the agent's intrinsic motivation for the work and I is an indicator function which is $I = 1$ if the agent is intrinsically motivated and $I = 0$ if not.[24]

Our incentive schemes define $b(y)$, the payment. It simplifies to

$$b(y) = pr \times y$$

for the two piecerate treatments with pr denoting the piecerate (either €0.02 or €0.1). In the two bonus treatments g denotes the target (either 50 or 100) and it can be written as

$$b(y) = \begin{cases} g \times 0.1, & \text{if } y \geq g \\ 0, & \text{if } y < g \end{cases}$$

The effort costs are represented by $c(e,i)$, which includes the explicit as well as implicit effort costs. The parameter i increases the marginal effort costs depending on the outside options available to the agent.[25] We assume that $c'_e(e, i_{Fix}) < c'_e(e, i_{Inet}) \leq c'_e(e, i_{Free})$ $\forall e \in [0, E]$, i.e. that marginal effort costs are higher in both environments, which provide outside options or alternative activities compared to the environment where subjects have to stay in front of the computer. Additionally, we assume the regularity conditions $\frac{\partial c(e,i)}{\partial e} > 0$ and $\frac{\partial^2 c(e,i)}{\partial e^2} > 0$ on the intervall $[0, E]$ and for simplicity, we assume $c(0,i) = 0$ $\forall i \in \{i_{Fix}, i_{Inet}, i_{Free}\}$. Furthermore, we assume that there is an effort level $E > 0$ at which effort costs increase to infinity, for example due to physical or time constraints, i.e. $\lim\limits_{e \to E} \frac{\partial c(e,i)}{\partial e} = \infty$.[26]

Piecerate Incentives:
We first discuss the two piecerate incentive schemes. Under those incentive schemes the maximization problems leads to the following first order condition. Agents supply effort as long as the marginal benefit of effort is higher than the marginal cost of effort.

$$\frac{\partial b(y)}{\partial y} \times \frac{\partial f(e)}{\partial e} + I\frac{\partial \delta(y)}{\partial y} \times \frac{\partial f(e)}{\partial e} = \frac{\partial c(e,i)}{\partial e}$$

[24]We assume $\delta'(y) \geq 0$.

[25]This includes for example versions of $c(e,i)$ as in Koch and Nafziger (2016) where $i \times \tilde{c}(e)$ and the parameter i differs between work environments, i.e. effort costs increase when alternative actions are present. Therefore $i_{Fix} \leq i_{Surf} \leq i_{Leave}$. This could also incorporate a version of $c(e,i)$, where implicit costs are modeled as utility of leisure, but leisure is negatively related with effort, i.e. time, as in Corgnet *et al.* (2015c).

[26]This is the similar to arguing that there is a maximal effort level subjects can exert in the experiment.

Let us now consider the difference of effort between work environments. We assume that marginal effort costs are higher in the environments with outside options or alternative activities. Therefore, with our work environment manipulation we increase the marginal effort costs in INET and FREE compared to FIX.

If we keep the incentive scheme as well as intrinsic and extrinsic marginal incentives constant effort changes only via a change in the marginal costs. It is easy to see that both work environments FREE and INET increase the marginal costs of effort and therefore decrease e^*, the optimal effort level and output decreases.

If we now compare the two piecerate incentive schemes, within a work environment, we only change the marginal benefits of effort. The marginal benefit equals the piecerate pr, which is larger in PIECERATE-HIGH than in PIECERATE-LOW. Therefore, the optimal effort level, i.e. output, increases in the piece-rate. However, it could be the case that subjects provide effort close to E and therefore output differences, i.e. differences in effort levels are negligible. Still, effort (i.e. output) in PIECERATE-HIGH should always be higher than in PIECERATE-LOW.

Bonus Incentives:
Bonus incentive schemes provide only marginal extrinsic incentives immediately at the target. However, they are not differentiable at that point. Therefore, we have to consider corner solutions and check the participation constraint. In the following, let $\hat{e}$ denote be the effort level, which is needed such that $g = f(\hat{e})$, i.e. the target is reached by the agent. We start by looking at the case without intrinsic motivation.

Case 1: Bonus incentives without intrinsic motivation
Without intrinsic motivation the maximization problem simplifies to

$$\max_{e \geq 0} u(e) = \bar{w} + b(y) - c(e, i).$$

Without intrinsic motivation it can never be optimal to exert effort $e \in (0, \hat{e})$, since the agent could always decrease effort and therefore his costs, without loosing any benefit. Similarly, it is easy to see that no effort above $\hat{e}$ can be optimal. Therefore the agent considers either to exert exactly the effort level $\hat{e}$, which is needed to reach the target and therefore the output equals $g = f(\hat{e})$, or to exert no effort at all. He exerts effort if the participation constraint is fulfilled, i.e.

$$\bar{w} + b(g) - c(\hat{e}, i) \geq \bar{w}.$$

Therefore, the agent exerts effort if reaching the target is beneficial for him, i.e. when the bonus payment is larger than the cost ($b(g) \geq c(\hat{e}, i)$).

Case 2: Bonus incentives with intrinsic motivation
With intrinsic motivation, additional solutions can arise. These solutions include points on the two intervals $[0, \hat{e})$ and $(\hat{e}, E]$. On these two intervals, the marginal benefits equal zero and therefore possible solutions have to fulfill the following condition.

$$I \frac{\partial \delta(y)}{\partial y} \times \frac{\partial f(e)}{\partial e} = \frac{\partial c(e, i)}{\partial e} \tag{1}$$

Which is equal to the first order condition, without marginal benefits. Let $\tilde{e}$ be the solution to this equation. This effort level $\tilde{e}$ can generally be above or below the effort level $\hat{e}$, which is needed to reach the target.

Case 2 a: Consider $\tilde{e} \geq \hat{e}$

If this is the case, $\tilde{e}$ is also an optimum, since monetary benefits are equal in both situations and providing effort above the target is optimal even in the absence of monetary incentives. Subjects exert effort until the marginal intrinsic motivation equals the marginal costs, which results in an even higher effort level as subject need in order to reach the target g.

This implies that for Bonus-Easy we would expect only few outputs above 50 as any additional output above the target would only be driven by workers which have a high intrinsic motivation. In Bonus-Hard subjects will not reach the threshold of 100 and therefore this case does not apply.

Case 2 b: Consider $\tilde{e} < \hat{e}$

If this is the case the agent has to check this local solution against the decision to exert an effort level $\hat{e}$, i.e. work until he reaches the target. Therefore he has to compare $\bar{w} + b(g) + I\delta(g) - c(\hat{e}, i)$ with $\bar{w} + I\delta(f(\tilde{e})) - c(\tilde{e}, i)$, where the exact solution depends on the exact form of the functions. The agent decides to exert effort level $\hat{e}$ if the additional costs of exerting the effort are lower than the additional benefits, i.e.

$$b(g) + I\delta(f(\hat{e})) - I\delta(f(\tilde{e})) \geq c(\hat{e}, i) - c(\tilde{e}, i)$$

Otherwise the agent will provide effort level $\tilde{e}$, which produces an output below the target. This case shows that there can be workers, which exert a positive effort level, which leads to an output below the target. Especially in the treatment Bonus-Hard subjects will not reach the threshold of 100. Therefore, observed output is due to workers, for which the intrinsic motivation induces the optimal effort $\tilde{e}$ such that $0 < \tilde{e} < \hat{e}$.

Let us now consider the difference of effort between work environments. Let us first consider *Case 1*. If the effort costs increase, due to a change in the work environment, it becomes more difficult to fulfill the participation constraint. Therefore some subjects will now exert less effort. For the *Case 2 a*, we can see that optimal effort decreases, if marginal effort costs increase. Since the intrinsic motivation does not change, higher marginal costs will induce lower effort. Also for *Case 2 b* effort can only decrease, if effort costs increase. Consider first those subjects which exert an effort level $\hat{e}$, some subjects might still exert effort until the target is reached. However, for some subjects it might be optimal to exert less effort, if effort costs increase. Those subjects, which already exerted an effort level below $\hat{e}$, will also decrease their effort.