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

Sebastian J. Goerg[‡] Sebastian Kube[§] Jonas Radbruch[¶]

This version: June 23, 2018

Abstract

Agents' decisions to exert effort depend 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 slider 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. In general, performance decreases as implicit costs increase. Yet, the magnitude of the reaction differs across incentive schemes and across the offered outside options; which, in turn, alters estimated output elasticities. In addition, comparisons between incentive schemes depend crucially on the implicit costs.

Keywords: Workers' Performance, Work Environments, Implicit Cost, Opportunity Costs, Incentive Schemes, Experiment

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

‡TU Munich, Max Planck Institute for Research on Collective Goods, and IZA, Schulgasse 22, 94315 Straubing, Germany. Email: s.goerg@tum.de

§University of Bonn, Max Planck Institute for Research on Collective Goods, and IZA, Adenauerallee 24–42, 53113 Bonn, Germany. Email: kube@uni-bonn.de

¶University of Bonn and IZA, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany. Email: radbruch@iza.org

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, behavior of agents 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 subjects work on the slider task (Gill and Prowse, 2012) while we manipulate opportunity costs and incentives.

In order 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 the real-effort slider 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 allowing 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 threshold is easy to achieve (Bonus-Easy), and in the second one it is (nearly) impossible to achieve (Bonus-Hard).

We observe higher output in the Fix-environment compared to the Inet- and Free-environments with increased implicit costs. Free results in an even sharper decrease as Inet. This result shows

[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.

that the increase in implicit effort costs decreases performance. The decrease of output compared to Fix can be observed across all incentive schemes, yet with different magnitudes for each incentive scheme. Free does not result in lower output for both piecerates compared to Inet, but does for both bonus-based incentive schemes. For the latter, the opportunity to leave the lab leads to a stronger decrease in output than the opportunity to use the internet. The different reaction to the introduction of implicit costs across incentive schemes and across implicit effort costs leads to differences in the comparison of incentive schemes, depending on the work environment. In the Fix-environment, all four incentive schemes result in rather similar outputs, although marginal incentives vary substantially. Only the high piecerate leads to a slightly higher output. In the Inet- and Free-environments, subjects are more likely to actually respond to incentives and we observe positive output elasticities for the response to piecerates.

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). The seminal work of Nalbantian and Schotter (1997), as well as many follow-up studies, examined 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). For example, people might show a negative response to the introduction of a very small piecerate (e.g., Gneezy and Rustichini, 2000), performance decreases as incentives become too large (e.g., Ariely *et al.*, 2009), or the strength of incentives and performances might generally follow an inverse u-shaped relationship (e.g., Pokorny, 2008). We demonstrate that not only the incentive side of the problem has to be taken into account, but that the opportunity cost side also plays a crucial part which is often neglected.

Methodologically, our paper adds to the literature using real-effort experiments, which are *"considered to be a better match to the field environment."* (Charness and Kuhn, 2011). Just recently, Herbst and Mas (2015) concluded in a meta-study on peer-effects that particularly experiments with real-effort tasks *"simulate realistic work environments"*. Real-effort 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 and Falk, 2011). So far, most of the experimental literature using real-effort experiments has considered fixed-time environments or fixed work requirements.[2] By the nature of those experiments, performance changes can only be due to a change in the explicit costs of effort.[3] One recent example of a study that changes the explicit effort cost is by Gächter *et al.* (2016), who combine a real-effort task with induced effort costs. In their study, the explicit costs of effort are exogenously varied by inducing

[2]One notable exception to this is Noussair and Stoop (2015), who use the time spent in the laboratory as a medium for reward. There, higher payoffs lead to shorter time in the lab. Similarly, Danilov and Vogelsang (2016) study time investments as pro-social giving.

[3]See Kurzban *et al.* (2013) for a related discussion in Psychology on salience and the effect of opportunity costs on task performance.

different costs for an action. Implicit costs play only a minor role in those experimental procedures, as subjects have to stay in the lab for a fixed time or until a task is completed. Other studies induce implicit costs through outside options, but do not vary them between treatments. Commonly used outside options are leaving the lab (e.g., Abeler *et al.*, 2011; Rosaz *et al.*, 2016), paid pause buttons (e.g., Mohnen *et al.*, 2008), surfing the internet (e.g., Corgnet *et al.*, 2015a), or reading magazines (e.g., Charness *et al.*, 2014).[4] However, those studies offer the outside option to every subject and do not manipulate the option.[5]

Our experiment is complemented by other studies that manipulate outside options (see e.g., Dickinson, 1999; Eckartz, 2014; Corgnet *et al.*, 2015c; Koch and Nafziger, 2016; Erkal *et al.*, 2018). The paper closest to ours is by Corgnet *et al.* (2015c). They study the effect of piecerate and team incentives while varying the access to one real-leisure option, namely internet browsing. Their key finding is that the availability of the real-leisure alternative leads to a sharper decrease in performances under team-based incentives than under piecerate incentives. Their study shows that implicit effort costs might play a role in determining effort. Our study takes this as a starting point to further investigate the role of implicit effort costs. The focus of our study, however, differs from their paper in at least two crucial aspects. First, the focus of our paper is on individual incentives studying two piecerate and two bonus schemes. Second, we manipulate the implicit costs in various ways and demonstrate that the effectiveness of the incentive schemes differs between work environments. Our study therefore investigates, in an unified framework, four commonly used individual incentives schemes in various environments. Our results demonstrate that the effectiveness of incentive schemes crucially depends 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 monetary incentive side of the problem, but simply due to the absence of implicit effort costs or to low opportunity costs. In addition, our study helps to explain why incentives sometimes might not change behavior in real-effort experiments (e.g., Araujo *et al.*, 2016). If individuals face (nearly) no costs for their effort or behavior, the corresponding behavior might not be altered by the incentive structure. Managers who are able to control the opportunity costs of effort directly might want to 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 should take into account that the behavioral responses to the incentive schemes depend 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.

[4] For more examples of papers using these outside options, see Tables 1 and 2 in the Online Appendix A.

[5] Some studies require outside options as only in their presence subjects are able to respond to the treatments, i.e., have a labor-leisure tradeoff (e.g., Kessler and Norton, 2016). In other studies, the usage of the outside option is the dependent variable of the experiment (e.g., Rosaz *et al.*, 2016).

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).[6] Each screen had 5 sliders that needed to be adjusted in order to finish a screen. The current number of finished screens was displayed on the screen (see Figure 3 in the Appendix for a screenshot) and the total number was later used to calculate the payments. The task was constant in all treatments and the effort to move the slider represented the explicit cost part of the opportunity costs. The experiment consisted of 3 stages and implicit effort costs were manipulated in the second stage.[7]

In the first stage, subjects worked on the real-effort task for 5 minutes without any monetary incentives. This stage served two purposes: First, subjects learned the difficulty of the task and could form accurate expectations about the effort costs, and secondly, it provided an ability measure which is not influenced by the subsequent incentive scheme.[8,9] Afterwards subjects received treatment-specific instructions and were informed about the subsequently applied incentives. Independent of the treatment, all instructions stressed that subjects should accomplish as many screens as possible.[10] The dependent variable — output, i.e., number of completed screens — was obtained in the second 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 third stage, subjects had to answer a short questionnaire, including sociodemographics, 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 in the second stage 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 piecerate 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.[11] In Bonus-Easy, subjects received a €5 bonus if

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

[7] An English translation of the instructions is provided in Online Appendix F.

[8] Strictly speaking, our ability measure is not able to differentiate between ability and intrinsic motivation. We thank an anonymous referee for pointing this out.

[9] Another possibility would have been to use an incentivized measure of ability. We choose not to incentivize this, since we didn't want subjects to experience different incentive schemes in the experiment.

[10] This was done as to minimize potential differences in crowding out of intrinsic motivation between work environments.

[11] See Gill *et al.* (2013) for a previous real-effort slider experiment with fixed targets. The targets in our experiment were chosen based on a pilot session with the real-effort task. The session had the same structure as the treatment Piecerate-High Fix.

Table 1: Treatments and Description of the 4×3-Design

Treatment Name	Incentives Scheme	Work Environment
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

they reached the target of 50 screens. This is a relatively easy target that most subjects could, and in fact did, reach. In Bonus-Hard, subjects received a bonus of €10 if they reached the target of 100 screens. This target was deliberately set very high and only one subject managed to reach the target.[12] The size of the bonuses were chosen such that they equated the earnings of a subject in the high piecerate 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 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 fixed-time procedure, in which subjects had to stay at the computer for 40 minutes without any leisure alternatives offered.[13] We manipulated the implicit costs by implementing two environments with alternative activities for the subjects. In the Inet-environment, subjects were allowed to use a web browser during the working phase of the experiment. Subjects had to remain in the laboratory for the whole time, but could surf the Internet instead of working on the task. This was implemented with a button on the real-effort screen, which would open a web browser and hide the real-effort task. Subjects could not work on the real-effort task and surf the Internet at the same time. However, they could always close the web browser and press a button to return to the real-effort task.[14] This allows us to record how

[12] As the target is set deliberately high, intrinsic motivation is the main driver of effort provision in this treatment (see also the derivation of the hypotheses in 2.1). Another possible explanation could be overconfidence. However, this would require a substantial amount of uncertainty about one's own capabilities – which seems unlikely in our setup, since all subjects should have been able to learn about the difficulty of the task during the first stage of the experiment.

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

[14] More details on the implementation of Inet and Free can be found in Online Appendix G.

much time subjects spent on the real-effort task and 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 led to the questionnaire and subjects could then leave the cubicle to get their payments. Payments were made based on the number of finished screens at the time the subject stopped working.

The experiments were conducted at the BonnEconLab of the University of Bonn. They were implemented using z-Tree (Fischbacher, 2007) 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 them from observing anything outside their cubicle. We conducted 16 regular sessions with the FIX- and INET-treatments and slightly adjusted the implementation in the FREE-treatments to prevent possible spillovers. In the FREE-treatments, subjects were invited to the lab on a given day, but could show up at any time between 10am and 4pm. This procedure ensured that subjects would not know the duration other subjects spend working on the task. In all treatments subjects received their payments individually in a separate room.

For each treatment, we gathered approximately 48 independent observations. In total, 571 subjects participated, with 58.6% of subjects being female and an average age of 23.58 years.[15] A session lasted on average 75 minutes for FIX and INET and individual sessions in FREE lasted 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 a total of €12.67, including the show-up fee. Between treatments, earnings ranged from €10 in the BONUS-HARD-treatments to a maximum of €20.5 in the PIECERATE-HIGH FIX treatment.[16]

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), intrinsic motivation $I\delta(y)$, and some costs depending on the explicit costs of effort e and the implicit costs i.[17] Implicit effort costs

[15]Neither age nor ability, as measured in the first stage, differ significantly between the three work environments ($p = 0.41$ and $p = 0.93$, both Kruskal-Wallis test). The gender composition differs slightly between treatments ($p = 0.099$, Kruskal-Wallis test). We use controls for gender, as well as age and ability, in our regression analyses to account for this.

[16]Table 6 in the Appendix presents implicit hourly wages per treatment.

[17]For example, this includes versions of $c(e, i)$ like in Koch and Nafziger (2016), where $c(e, i) = i \cdot \tilde{c}(e)$, and where the parameter i differs between work environments, i.e., effort costs increase when alternative actions are present. Therefore $i_{Fix} \leq i_{Inet} \leq i_{Free}$. 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)

in our setup represent the foregone utility of not allocating the effort or time to other activities. Following the approaches by Murdock (2002) and James (2005), δ represents the agent's intrinsic motivation for the work (if she is intrinsically motivated) and I is an indicator function which is $I = 1$ if the agent is intrinsically motivated or $I = 0$ if not. In what follows, we present the intuition underlying our behavioral predictions and discuss the framework in more detail in the Online Appendix B.

With our work environment manipulation, which changes the implicit costs, we increase the marginal effort costs in INET and FREE compared to FIX. Furthermore, it seems reasonable to assume that subjects in FREE have more outside opportunities than in INET. This would imply an additional increase of the marginal effort costs in FREE compared to INET. Output should decrease as the marginal costs of effort increases. Thus, we expect the highest output in FIX (since implicit costs are low) and the lowest output in FREE (since implicit costs are high).

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

Let us now consider the differences between the piecerate treatments. Subjects provide effort as long as the marginal benefits from the piecerate payment and the intrinsic motivation to perform the task are higher than 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 piecerate treatments within a work environment.

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

For BONUS-EASY we would expect only few outputs above 50 as $\frac{\partial b(y)}{\partial y} = 0$ for any output above 50. Additional output would only be driven by workers for whom 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. Consequently, output would again only be driven by workers for whom the marginal intrinsic motivation is higher than the marginal costs.

The differences between BONUS-EASY and the two piecerate treatments are ultimately an empirical question, because predictions about performance differences would require additional assumptions about the exact form of the cost of effort function and the intrinsic motivation. However, since reaching the target of 100 screens in BONUS-HARD is not feasible, we can predict that in both piecerate treatments output should be higher than in BONUS-HARD. This is due to the fact that monetary incentives are basically absent in BONUS-HARD and therefore incentives are higher in the two piecerate treatments.

Hypothesis 3: *We expect higher output in* Bonus-Easy *than in* Bonus-Hard*. The output in* Bonus-Easy *should be 50 screens or slightly above. Furthermore, we expect higher output in both piecerate treatments than in* Bonus-Hard*.*

3 Results

Table 2 provides summary statistics for the output in all treatments. In the following, we will first demonstrate that implicit costs have a significant impact on work output and discuss their influence within an incentive scheme. Thereafter, we will demonstrate that implicit costs influence the comparisons between incentive contracts. Finally, we will take a closer look at the usage of the offered outside option and the influence of non-cognitive traits on behavior. If not stated otherwise, reported p-values are two-sided and based on t-tests and regressions.[18]

Table 2: Summary Statistics of Outputs

		Over All	By Incentive Scheme			
		Incentives	Piecerate-Low	Piecerate-High	Bonus-Easy	Bonus-Hard
Fix	Mean	58.79	57.56	59.40	60.09	58.21
	SD	14.44	14.59	13.59	12.49	16.91
	N	188	48	47	45	48
Inet	Mean	50.82	41.87	53.69	53.19	54.43
	SD	19.90	23.47	16.50	15.50	21.06
	N	190	47	48	48	47
Free	Mean	43.34	36.21	52.35	45.94	38.69
	SD	24.42	27.18	21.02	20.41	25.79
	N	193	48	49	48	48

SD: standard deviation, N: number of independent observations

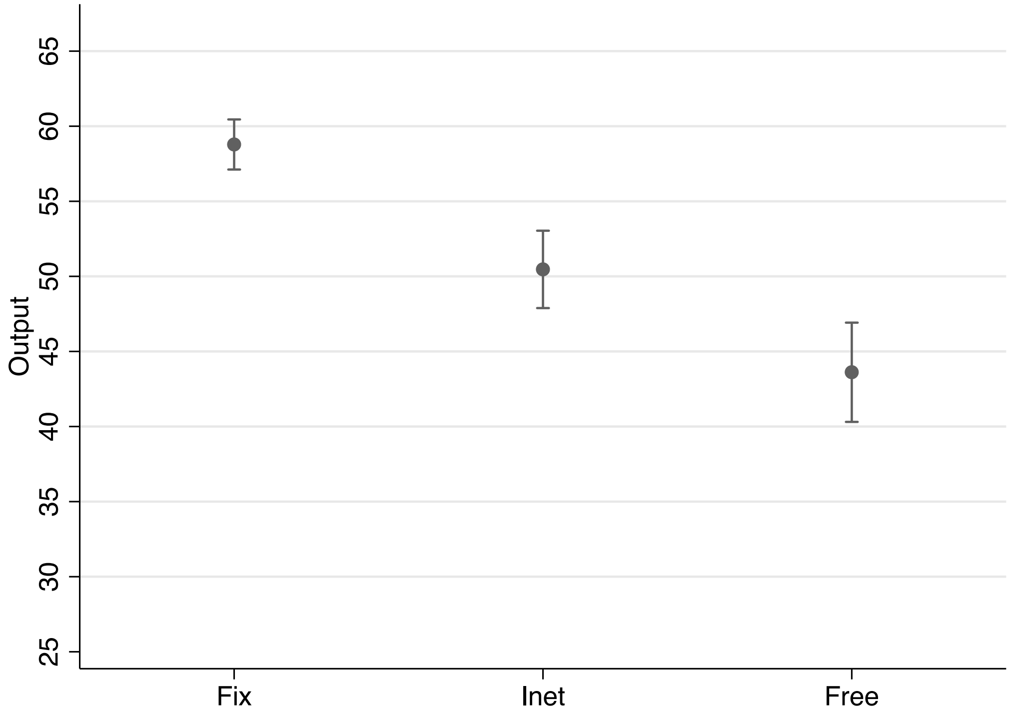
3.1 The impact of implicit costs on output

We start by looking at the general effect of implicit costs for all incentive schemes. Based on the raw means reported in Table 2, output in the Fix-treatments is on average 15.7% higher than in the Inet-treatments and 35.6% higher than in the Free-treatments. The average output in the Inet-treatments is 17.3% higher than in the Free-treatments. The predicted output of our three work environments is presented in Figure 1. The figure is based on a least squares regression with controls for ability, gender, and age.[19] In general, Figure 1 shows that output decreases significantly with higher implicit costs (all pairwise comparisons $p < 0.01$). Already based on this general look

[18] Additionally, Tables 9 and 10 in the Appendix report p-values of non-parametric tests.

[19] The corresponding regression table is presented in Table 7 in the Appendix. Furthermore, Figure 4 in the Appendix displays the boxplots for the output level for each incentive scheme.

Figure 1: Predicted Output with 95% CIs



Estimates are based on linear regression controlling for subjects' ability, gender, and age. Plot shows the margins with confidence intervals. For results and coefficients of the corresponding regressions, see Table 7 in the Appendix.

at the data we can conclude that implicit costs in general influence the output negatively, which is in line with our predictions.

However, the impact of implicit costs is not limited to the average outputs; Table 2 and Figure 1 reveal that implicit costs increase the variance of the output, too. The variance differs significantly between work environments and increases with opportunity costs (all $p < 0.01$, using two-sided Variance-ratio tests). The lowest variance is observed in Fix, increases in Inet, and is highest in Free.[20] Our treatments increase the implicit costs by manipulating outside options and the time spent on the outside option reduces the output. However, not all subjects utilize the outside options to the same extent, which increases the variance. In fact, the largest part of the observed variance is explained by the total time spent working on the task (see Section 3.4).[21]

Result 1: *In line with Hypothesis 1, subjects' output decreases significantly as implicit effort costs increase. At the same time, the variance of output increases with implicit costs.*

[20]This pattern generally holds for each incentive scheme individually.

[21]A more detailed analysis of the time worked on the task can be found in Online Appendix D. Table 6 in Online Appendix D demonstrates that the time subjects work on the task changes analogously to the changes in output presented here.

Table 3: Regression of Output on Work Environments

	Piecerate-Low		Piecerate-High		Bonus-Easy		Bonus-Hard	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Inet	-15.69***	-12.55***	-5.72*	-6.15**	-6.90**	-7.41***	-3.78	-3.85
	(4.02)	(3.37)	(3.10)	(2.65)	(2.91)	(2.47)	(3.92)	(3.43)
Free	-21.35***	-17.28***	-7.06*	-8.23***	-14.15***	-14.87***	-19.52***	-17.23***
	(4.45)	(4.25)	(3.60)	(3.01)	(3.48)	(3.37)	(4.45)	(4.42)
Constant	57.56***	1.30	59.40***	35.10***	60.09***	47.44***	58.21***	28.62**
	(2.11)	(12.28)	(1.98)	(6.99)	(1.86)	(13.04)	(2.44)	(11.14)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	143	142	144	142	141	141	143	143
R^2	.14	.35	.031	.31	.11	.23	.14	.26
(Inet and Free = 0)	0.00	0.00	0.07	0.01	0.00	0.00	0.00	0.00
(Inet = Free)	0.28	0.33	0.73	0.55	0.05	0.04	0.00	0.00

Table presents least squares regression using performance as dependent variable. * p < 0.1, ** p < 0.05, *** p < 0.01. Robust Standard errors in parentheses. For some observations controls are missing, due to some subjects who refused to answer some of the sociodemographic questions. Controls: ability, age, gender.

In the following, we turn to differences within each incentive scheme. Table 3 estimates the treatment effects for each incentive scheme using Fix as the benchmark category. In Piecerate-Low, we observe a decline of effort between Fix and both Inet and Free. Output decreases when subjects face increased implicit effort costs, compared to Fix. With added controls, average output significantly decreases by 12.55 screens in Inet and by 17.28 screens in Free (both coefficients with $p < 0.01$). Average output in Piecerate-Low Free is lower than in Piecerate-Low Inet, but this difference turns out to be insignificant (p=0.28).

A similar pattern emerges for Piecerate-High. Compared to Fix, output decreases by 6.15 screens in Inet ($p < 0.05$) and by 8.23 screens in Free ($p < 0.01$). However, the two coefficients are smaller than their counterparts in Piecerate-Low. Again, output is the lowest in Free, but Fix and Inet do not differ significantly from each other ($p = 0.55$). Our results are therefore in line with the first part of Hypothesis 1, but not with the second part that Free induces higher implicit costs than Inet.[22]

Result 2: *Implicit costs significantly reduce the performance in the incentive schemes* Piecerate-High *and* Piecerate-Low*. Differences between* Inet *and* Free *exist, but do not turn out to be significant.*

Similar to the two piecerate 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,

[22] In the Online Appendix C, we parameterize our theoretical framework and estimate a structural model. Table 3 in the Online Appendix C presents the results of this exercise. The estimation results support the notion that the presence of outside options change the implicit effort costs. Depending on the exact parametrization, the marginal implicit effort costs are estimated between 8.4 and 17.8 cents for Inet and between 10.1 and 18.3 cents for Free.

the output decreases by 7.41 screens in INET ($p < 0.01$) and 14.87 screens in FREE ($p < 0.01$) compared to FIX. The output in BONUS-EASY INET is 15.6 % higher compared to FREE and is also significantly different ($p = 0.04$). As Figure 4 in the Appendix shows, the output distribution in BONUS-EASY FREE collapses around 50 screens. In line with Hypothesis 3, the majority of subjects stop working once they reached the threshold for the bonus.[23] Only 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 BONUS-EASY FIX, 88.89% of subjects reached the target of 50 screens; in INET, 85.42%; and in FREE, 77.08% (all pairwise comparisons $p > 0.108$ or above, two-sided Fisher's exact test).

In BONUS-HARD, we observe the same pattern in output levels. In BONUS-HARD INET, the average output does not differ significantly from the average output in FIX. In BONUS-HARD FREE, average output is 17.23 screens lower than in FIX ($p < 0.01$) and 13.38 screens lower than in INET ($p < 0.01$). The output distribution in BONUS-HARD FREE is shifted downwards and has a longer lower tail (again, compare Figure 4 in the Appendix). This is mostly driven by subjects who stop working and leave the lab early. Only one subject in the BONUS-HARD-treatments was able to reach the target of 100 screens.[24]

Our results show that for both bonus based incentive schemes output decreases in FREE compared to INET. This is in line with our hypothesis that implicit costs are increasing in FREE resulting in lower output. Interestingly, more subjects work very short times on the task and produce very low outputs in FREE. In fact, in BONUS-EASY, the number of subjects who work less than 5 minutes increases from one in INET to seven in FREE. Similarly, in BONUS-HARD this number increases from one to seven. Thus, more subjects exert very low levels of effort.[25]

Result 3: *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. *Unlike the two piecerate treatments, outputs differ significantly between* INET *and* FREE *in both bonus treatments.*

[23] Again, the different outputs result from different durations spent working on the task. Refer to Online Appendix–D for an additional analysis.

[24] We use the estimated parameters of the structural model in the Online Appendix C to derive predictions about the two discretionary bonus treatments (see Table 4 in Online Appendix C). Again, the results of the structural model support the analyses presented in this section.

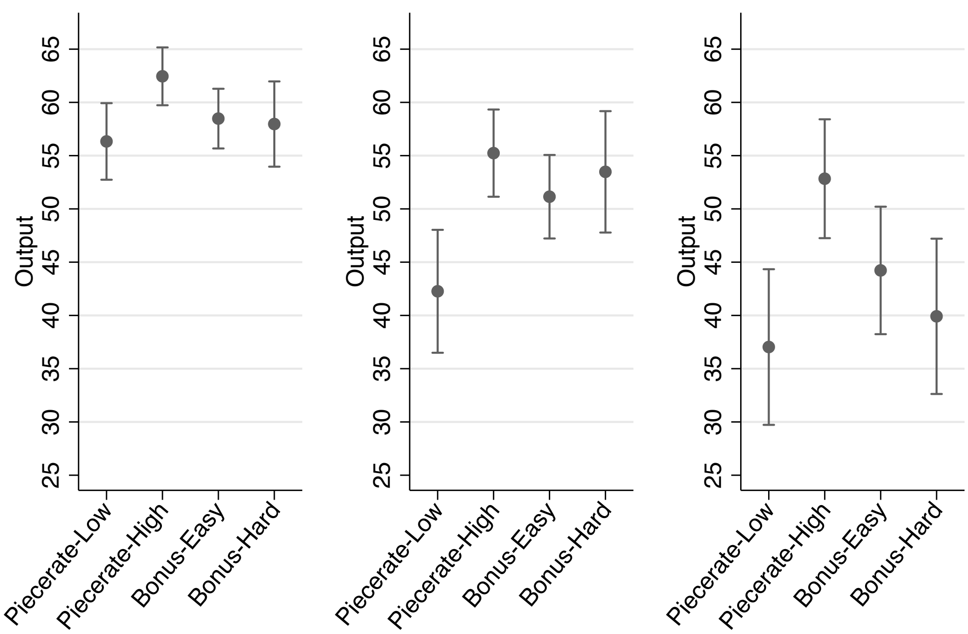
[25] These findings are also reflected in subjects' outputs. In BONUS-EASY INET (BONUS-HARD INET) two (one) subjects have an output below ten; this number increases to seven (nine) in BONUS-EASY FREE (BONUS-HARD FREE). One possible explanation is that increasing costs are more likely to trigger the theoretical corner solution of the two bonus treatments.

3.2 Implicit costs and the comparison between incentive schemes

So far we have demonstrated that implicit costs can influence the output even if marginal monetary incentives are fixed. In a next step, we will demonstrate that implicit costs influence the comparison of incentive schemes. The descriptive statistics are provided in Table 2, while Figure 2 presents the estimated output in all treatments after controlling for ability, gender, and age. The figure is based on the estimation results presented in Table 8 of the Appendix.[26]

Figure 2: PREDICTED OUTPUT WITH 95% CIS

(a) FIX (b) INET (c) FREE



Estimates are based on linear regressions controlling for subjects' ability, gender, and age. Plot shows the margins with confidence intervals. For results and coefficients of the corresponding regressions, see Table 8 in the Appendix.

In the FIX-treatments, the highest output is observed in PIECERATE-HIGH and the lowest output in PIECERATE-LOW.[27] Output in PIECERATE-HIGH is significantly higher than in PIECERATE-LOW ($p = 0.012$), BONUS-EASY ($p = 0.048$), and BONUS-HARD ($p = 0.078$).[28] All other comparisons are insignificant ($p \geq 0.35$). Comparing output across incentive schemes for INET, we observe again the highest output in PIECERATE-HIGH and the lowest in PIECERATE-LOW. Output in PIECERATE-

[26] We estimate one regression per work environment controlling for ability, age and gender. All reported p-values in this part of the analysis are based on these regressions.

[27] It is worth pointing out that without additional controls we fail to identify any significant differences between the treatments in FIX.

[28] All p-values in this part of the analysis are based on the regressions in Table 8 in the Appendix using the specifications with control variables. Figure 2 is based on the same regression table.

Low turns out to be significantly lower than in the other three incentive schemes (all $p < 0.001$). Yet, all other comparisons remain insignificant ($p \geq 0.17$). If we compare the incentive schemes within the working environment Free, we again observe the highest output in Piecerate-High and the lowest in Piecerate-Low. In Free, output in Piecerate-High is again significantly higher than in Piecerate-Low ($p = 0.001$), Bonus-Easy ($p = 0.04$), and Bonus-Hard ($p = 0.005$). Outputs in Piecerate-Low, Bonus-Easy, and Bonus-Hard do not differ significantly ($p > .12$ for all pairwise comparisons).

Comparing these reported differences across work environments provides additional insights into the impact of implicit costs. Increasing the implicit costs, we observe stronger negative reactions for Piecerate-Low than for Piecerate-High. This influences the comparison between the two incentive schemes. In Fix, average output in Piecerate-High is only 11% higher than in Piecerate-Low; in Inet it is 31% higher; and in Free, it is even 43% higher.[29] In both Inet and Free, these changes are significantly larger than in Fix (both $p < 0.05$).

Output in the two bonus schemes responds to increased implicit costs. However, implicit costs do not significantly influence the comparison between Bonus-Easy and Bonus-Hard. In all three environments, Fix, Inet, and Free, we observe no significant differences between the outputs in Bonus-Easy and Bonus-Hard (all $p \geq 0.38$). However, comparing bonus schemes with piecerates shows that the outputs respond differently to changes in implicit costs (see Figure 2).

Average output under Piecerate-Low decreases most strongly with the introduction of implicit costs (moving from Fix to Inet), while in the bonus schemes the response tends to be stronger if implicit costs increase further (moving from Inet to Free). This influences the comparison between the piecerate schemes and the bonus schemes. Without implicit costs, output does not differ significantly between Piecerate-Low, Bonus-Easy, and Bonus-Hard (all pairwise comparisons $p \geq 0.37$). However, the steep decline in Piecerate-Low Inet results in significantly lower output compared to Bonus-Easy Inet ($p < 0.01$) and Bonus-Hard Inet ($p < 0.01$). Yet, after the output declines more steeply in Bonus-Easy Free and Bonus-Hard Free, outputs are no longer significantly different between the two bonus schemes and Piecerate-Low Free (both $p \geq .13$).

With the introduction of implicit costs, average output also decreases under Piecerate-High, but not as strongly as under Piecerate-Low. At the same time, increased implicit costs increase the variance. Thus, we observe quite the opposite picture when comparing Piecerate-High with Bonus-Easy and Bonus-Hard. While in Fix output is significantly higher in Piecerate-High than in Bonus-Easy and Bonus-Hard (both $p < 0.05$), they no longer differ significantly in Inet. Only after average output in the two bonus schemes declines steeply in Free, is output in Piecerate-High significantly higher than output in Bonus-Easy ($p < 0.05$) and Bonus-Hard

[29]Percentages are based on the predictive margins presented in Figure 2, which are based on the regressions in Table 8.

($p < 0.01$). To summarize, we find partial support for Hypothesis 3, but output in Bonus-Hard tends to be higher than expected.

Result 4: *Of all incentive schemes,* Piecerate-High *responds least to changes in the implicit cost; in contrast we observe a strong response in* Piecerate-Low. *Therefore, the difference in outputs between the two piecerate schemes increases with implicit costs. The comparison between the bonus- and piecerate-schemes depends on the exact setting.*

3.3 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 pure incentive theory perspective, we would on average expect higher outputs in Piecerate-High than in Piecerate-Low, i.e., a positive output elasticity. Table 4 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 that turns significant only after adding controls. For both, Inet and Free, we observe significantly larger and positive elasticities compared to Fix ($p = 0.0287$ and $p < 0.01$, two-sided, Wald test). Increasing the piecerate by 1% would increase the outputs by 0.25% in Inet and by 0.52% in Free. However, the difference between these two elasticities falls short of reaching conventional levels of significance ($p = 0.107$).

Table 4: Elasticities

	Fix		Inet		Free	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Piecerate)	0.0307	0.0841*	0.2577***	0.2722***	0.5272***	0.5231***
	(0.0386)	(0.0438)	(0.0973)	(0.0896)	(0.1378)	(0.1352)
Constant	4.1269***	3.5550***	4.5126***	3.8518***	5.0825***	3.1893***
	(0.1136)	(0.1401)	(0.2518)	(0.6359)	(0.3539)	(0.5967)
Controls	No	Yes	No	Yes	No	Yes
N	95	95	93	92	95	93
R^2	.0067	.34	.075	.25	.14	.23

Table presents least squares regression using the logarithm of performance as dependent variable. * p < 0.1, ** p < 0.05, *** p < 0.01. Robust Standard errors in parentheses. Observations with a performance of zero are dropped from the estimation. Controls: ability, age, gender.

These results show that implicit costs induced by different work environments matter. While subjects' output only responds marginally to increased incentives in Fix, we are able to observe significant reactions in outputs to increased incentives in both work environments with higher implicit costs. While the response in Inet and Free is positive, it is still inelastic.

Result 5: *The elasticity of the output increases with implicit costs. In* Fix*, the elasticity of the output differs only weakly significantly from zero. By contrast, in both* Inet *and* Free *we observe a positive and significant response of the output to increased incentives.*

3.4 Supplementary Analyses

We observe a high variance in performance across all incentive schemes and work environments, especially in those with an outside option, i.e., Inet and Free. Therefore, in the following we take a closer look at individual characteristics and personality traits and their impact in the different work environments. This will help us to understand to what extent the observed variance is driven by those characteristics. We therefore regress output on characteristics interacted with an indicator for each environment in Table 5. All reported models include treatment fixed effects.

Table 5: Determinants of output

	(1)	(2)	(3)	(4)	(5)
Fix x Ability	6.1241***			6.3441***	6.3441***
	(0.4908)			(0.5681)	(0.5686)
Inet x Ability	5.1249***			5.1377***	4.2718***
	(0.8575)			(0.8900)	(0.5149)
Free x Ability	4.1733***			4.4898***	3.9106***
	(1.0097)			(1.0956)	(0.4158)
Fix x CRT score		1.1676		-0.1180	-0.1180
		(0.8405)		(0.6849)	(0.6855)
Inet x CRT score		2.6692**		1.7756	0.8935
		(1.2194)		(1.1584)	(0.5818)
Free x CRT score		-0.8770		-2.1454	0.4436
		(1.5251)		(1.6693)	(0.6526)
Fix x Conscientiousness			-0.4496	0.0290	0.0290
			(0.6417)	(0.4674)	(0.4679)
Inet x Conscientiousness			1.3723*	1.1101	-0.3025
			(0.8124)	(0.7738)	(0.4219)
Free x Conscientiousness			0.6093	0.0464	-0.2763
			(0.9522)	(1.0703)	(0.4275)
Total time					1.5185***
					(0.0325)
Treatment FE	Yes	Yes	Yes	Yes	Yes
Gender and Age	No	No	No	Yes	Yes
Big 5 (w/o Cons), Risk	No	No	No	Yes	Yes
N	569	571	571	568	568
R^2	.29	.15	.15	.32	.8

Table presents least squares regression using output as dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust Standard errors in parentheses. Big 5 (w/o Cons.) controls for the other Big 5 traits and risk for general risk attitudes all interacted with a variable indicating the work environment.

In all implicit cost settings, the output and the ability measures have a significant positive relationship, but the strength differs (see Model 1). For each finished screen in the ability stage, subjects are estimated to complete slightly more than 6 screens in the main experiment in the FIX-treatments, 5 screens in the INET-treatments, and 4 screens in the FREE-treatments. The influence of an agent's ability on output differs significantly between FIX and FREE (p-value = 0.082, two-sided, Wald-test), but not for any other comparison (all p-values> .31, two-sided, Wald-test).[30]

Furthermore, we can explore the relationship between personal characteristics and output. 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).[31] 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. For strategic interactions, Gill and Prowse (2017) find neither a correlation between response times and personality nor between response times and cognitive ability.

We find a positive relationship between the CRT score and output (Model 2) in INET, but not in the other environments. Similarly, conscientiousness (Model 3) is positively associated with output in INET, but not in the other environments. In Model 4, we simultaneously control for all measures. As several of the measures are correlated with each other, only the influence of ability remains significant. In a last step, we additionally include the total time (in minutes) worked on the task (Model 5). This increases the explained variance dramatically as the R^2 improves from .32 in Model 4 to .80 in Model 5. Obviously, the time subjects worked on the task explains the largest part of the variance in our data.[32]

In fact, the reported differences in outputs result from a substantial fraction of subjects using the outside options when available: 36.84% in INET and 47.67% in FREE. In both work environments, the usage of the outside option reduces the time subjects spent working on the task. In INET and FREE, the average working time is significantly below 40 minutes (both $p < 0.01$). Subjects work on the task on average 35.12 minutes in INET and only 29.18 minutes in FREE ($p < 0.01$). Thus,

[30] Recall that ability did not differ significantly between treatments.

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

[32] In the Online Appendix E, we also repeat the analysis from above and use questionnaire data to explore the influence of personality measures on the time subjects spent working on the task. The direction of the point estimates is in line with our results in Table 5.

our treatments influence more the extensive margin (time spent working on task) than the intensive margin (speed while working on task).[33]

We conclude our last result:

Result 6: *Conscientiousness and CRT are significantly correlated with higher output in* Inet. *The time subjects spent working on the task explains a large part of the observed variance in output. Subjects time on the task decreases significantly as implicit effort costs increase.*

4 Discussion and Conclusion

In this paper, we investigate how work environments with different implicit costs influence the effectiveness of linear and non-linear incentive schemes. We exogenously vary the implicit effort costs between work environments by offering real-leisure alternatives and comparing the performance of subjects in two piecerates and two bonus schemes. We observe that incentive contracts and opportunity costs interact in a non-trivial manner. Generally, as implicit costs increase, the average output decreases and the variance of output increases. Yet, the responses are not equally strong for all incentive schemes. We observe stronger negative reactions for Piecerate-Low than for Piecerate-High. These unequal reactions lead to increasing differences among those two incentive schemes: in Fix, average output under the high piecerate is 11% higher than under the low piecerate; in Inet, it is 31% higher; and in Free, even 43% higher. Likewise, an increase in implicit costs increases the output elasticity of piecerates. With respect to non-linear incentive schemes, our results suggest that the effect of bonus schemes depends on the opportunities of workers to allocate their time. Our results in Bonus-Easy suggest that achievable targets induce behavior such that targets are closely matched, but not exceeded, in those work environments with substantial implicit effort costs. For targets like Bonus-Hard, implicit costs increase the number of workers who drop out of the task once they realize that the target is hard to achieve. However, in the Fix-environment with low implicit costs, we observe, for both bonus schemes, an effort that is far from any incentivized points – either beyond the target or far before the target is reached. This behavior might be more in line with subjects who consider this a fixed wage setting than a bonus setting. Although this might be unexpected, it is similar to the fixed bonus treatments reported in DellaVigna and Pope (2018), where experts also fail to forecast the effort provision beyond an incentivized point. Moreover, monetary incentives differ strongly between Piecerate-High and Bonus-Hard; yet, only in Free we do observe a large and highly significant difference between the two incentive schemes. Thus, our results in general show the dependency of the effectiveness of incentive schemes with respect to the work environment, i.e., the implicit effort costs. For example, workers in bonus schemes might be less sensitive to incentives in environments similar to Fix and Inet. Behavior in those environments might not be well predicted by standard incentive theory.

[33] In the Online Appendix D, we also show that, once we account for the time subjects work on the task, output is similar across all environments.

However, as the implicit costs of effort increase, the behavior aligns more and more with predictions made by incentive theory.

In addition to providing new insights into the interplay of piecerates and bonus schemes with implicit effort costs, our paper also confirms and qualifies previous findings in the literature. In general, increasing the implicit costs of effort, while keeping the incentives (wages or piecerates) fixed, leads to smaller output (Corgnet *et al.*, 2015c; Koch and Nafziger, 2016). Yet, a superficial look might suggest that our results are not fully in line with Corgnet *et al.* (2015c). While Corgnet *et al.* (2015c) find no significant impact of the option to surf the Internet under a piecerate contract, we observe significant differences between Fix and Inet under two piecerate contracts. However, Corgnet *et al.* (2015c) also report a 10% smaller output if the option to surf the Internet is available. In our most comparable treatments, Piecerate-High Fix and Piecerate-High Inet, we observe the same drop of output by 10%.[34] Furthermore, their results, over time, demonstrate significant differences in the later part of their experiment. Figure 5 in the Appendix demonstrates that the comparison of output in Piecerate-High Fix and Piecerate-High Inet in our paper follows the same dynamics. For Piecerate-High, we initially observe no significant difference, but over time a pronounced difference develops between the outputs in Fix and Inet. Thus, we confirm the finding by Corgnet *et al.* (2015c) that implicit cost effects under high piecerates are dynamic and need some time to develop, although in our setting these effects are strong enough to result in overall significant differences. Furthermore, we demonstrate that the elasticity of the output increases with implicit effort costs. Moreover, our results show that implicit effort costs and the exact nature of those might be even more important for bonus-based incentive contracts. Due to the existence – and depending on the exact size – of implicit costs, subjects might either explicitly target the bonus or abstain from working completely. Our paper also contributes to the ongoing discussion of the real-effort slider task and real-effort experiments in general. Araujo *et al.* (2016) implement the slider task in a fixed laboratory environment with three different piecerate schemes and conclude that it demonstrates no meaningful response to explicit monetary incentives. We show that this is more a problem of the fixed laboratory environment than the slider task itself. Our estimated output elasticity of 0.0307 in Fix is very similar to the elasticity of 0.025 estimated by Araujo *et al.* (2016). Yet, once implicit costs are increased, subjects respond in a meaningful and significant way to the linear incentives. Thus, effort in any (real-effort) task should not be evaluated independently of the work environment it is implemented in.

Our results have implications for the use and design of incentive schemes within organizations. The management has many means to affect worker behavior and every aspect of an organization can be used as a parameter to obtain desired outcomes (Roberts, 2007). In addition to monetary and non-monetary incentives, organizations should recognize that they might want to adjust implicit costs as a relevant parameter, too. For example, to increase the output in our Piecerate-Low

[34]Comparability is based on the available outside option and implicit hourly wages. The implicit wage is calculated by using the performance-dependent pay component (payoff without show-up fee) and scaling it up/down to an hourly wage.

INET setting, one could either implement a higher piecerate (PIECERATE-HIGH INET) or keep the piecerate fixed and reduce the implicit costs (PIECERATE-LOW FIX). Using our experimental results, the first approach would, on average, increase output by roughly 29% with additional costs of € 4.56 per worker, and the second approach would increase the output by roughly 40% with additional costs of only € 0.34. Even if the management cannot change the work environment, it is important to take implicit costs into account when implementing and evaluating traditional incentive schemes. Our analysis of output elasticities would suggest that there are only minor benefits from increased piecerates in environments similar to our FIX-treatments, but larger gains from changes in the piecerate in environments similar to our FREE-treatments. Similarly to managers in firms, unemployment agencies want to incentivize job seekers to find a job.[35] Job seekers have to seek their jobs in an environment where leisure costs are potentially high and leisure alternatives are easily available and always present. Unemployment agencies therefore could use a FIX environment, for example by requiring job seekers to spend a fixed amount of time in a room with access to material needed for applications but no leisure alternatives. Beyond the analysis of incentive schemes, our results and implications are also interesting in light of the recent discussion of workplace flexibility and home offices. Given our results it is not surprising that, after the boom of telecommuting in the last decade, companies like IBM are now adopting more restrictive approaches to home office and telecommuting and either demand full presence or at least required presence times.[36] Other firms, for example call centers, incentivize their flexible workers to work specific hours, using contracts which yield bonuses for making calls for a given time in the evening hours.[37] Our paper demonstrates that reducing implicit costs and temptations like surfing the internet leads to higher productivity. Yet, some caution is warranted as workers might realize that the work environment is an active choice by the management and introduce reciprocal motives. As such, the active choice of inflexible work environments, which reduce implicit effort costs, might signal distrust and reduce motivation and output of the worker (Alder *et al.*, 2006; Corgnet *et al.*, 2015d; Koch and Nafziger, 2016). More generally, controlling the own work environment can influence workers motivation (Deci *et al.*, 1989; Deci and Ryan, 1995) and change the performance of individuals (Kiessling *et al.*, 2018).

Future work on implicit costs in work environments should extend to non-monotone tasks that require creativity, communication, and innovation. Apple, Unilever, and Facebook are just a few examples of firms that use architecture to design work environments encouraging communication and serendipitous encounters through coffee places and meeting points.[38] While these new work environments are intended to increase innovation and creativity, they also increase the implicit costs of effort. Investigating the net effect in such environments seems to be an important next step.

[35] Instead of relying on incentive contracts, unemployment agencies rely, for example, on binding job search requirements (e.g., Arni and Schiprowski, 2017).

[36] See https://www.bloomberg.com/news/articles/2017-07-10/the-rise-and-fall-of-working-from-home.

[37] One example is Infas in Bonn. Information is provided on their website: www.infas.de

[38] Accessible introductions to this topic are provided by Wagner and Watch (2017) and Waber *et al.* (2014).

References

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

Alder, G. S., Noel, T. W. and Ambrose, M. L. (2006). Clarifying the effects of internet monitoring on job attitudes: The mediating role of employee trust. *Information & Management*, **43** (7), 894 – 903.

Araujo, F. A., Carbone, E., Conell-Price, L., Dunietz, M. W., 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. *Journal of the Economic Science Association*, **2** (1), 1–12.

Ariely, D., Gneezy, U., Loewenstein, G. and Mazar, N. (2009). Large Stakes and Big Mistakes. *The Review of Economic Studies*, **76** (2), 451–469.

Arni, P. and Schiprowski, A. (2017). Job Search Requirements, Effort Provision and Labor Market Outcomes.

Asch, B. (1990). Do Incentives Matter? The Case of Navy Recruiters. *Industrial & Labor Relations Review*, **43** (3), 89S–106S.

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*, **62** (11), 3085–3099.

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. (2014). The Dark Side of Competition for Status. *Management Science*, **60** (1), 38–55.

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.

—, Hernán-González, R. and McCarter, M. W. (2015d). The role of the decision-making regime on cooperation in a workgroup social dilemma: An examination of cyberloafing. *Games*, **6** (4), 588–603.

Danilov, A. and Vogelsang, T. (2016). Time for Helping. *Journal of the Economic Science Association*, **2** (1), 36–47.

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

—, Connell, J. P. and Ryan, R. M. (1989). Self-determination in a work organization. *Journal of applied psychology*, **74** (4), 580.

— and Ryan, R. (1995). Intrinsic motivation and self-determination in human behavior. *New York*.

DellaVigna, S. and Pope, D. (2018). What Motivates Effort? Evidence and Expert Forecasts. *The Review of Economic Studies,*, **85** (2), 1029–1069.

Dickinson, D. L. (1999). An Experimental Examination of Labor Supply and Work Intensities. *Journal of Labor Economics*, **17** (4), 638–670.

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

—, —, 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.

Eckartz, K. (2014). Task Enjoyment and Opportunity Costs in the Lab: The Effect of Financial Incentives on Performance in Real Effort Tasks. *Jena Economic Research Papers*.

Erkal, N., Gangadharan, L. and Koh, B. H. (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review*, **101**, 528 – 545.

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*, **19** (4), 25–42.

Gächter, S., Huang, L. and Sefton, M. (2016). Combining "Real Effort" with Induced Effort Costs: the Ball-Catching Task. *Experimental Economics*, **19** (4), 687–712.

Gill, D. and Prowse, V. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *The American economic review*, **102** (1), 469–503.

—, — and Vlassopoulos, M. (2013). Cheating in the Workplace: An Experimental Study of the Impact of Bonuses and Productivity. *Journal of Economic Behavior & Organization*, **96**, 120–134.

— and Prowse, V. L. (2017). Using Response Times to Measure Strategic Complexity and the Value of Thinking in Games.

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.

— and Rustichini, A. (2000). Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, **115** (3), 791–810.

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.

Herbst, D. and Mas, A. (2015). Peer Effects on Worker Output in the Laboratory Generalize to the Field. *Science*, **350** (6260), 545–549.

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*, **55** (2), 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.

Kessler, J. B. and Norton, M. I. (2016). Tax Aversion in Labor Supply. *Journal of Economic Behavior and Organization*, **124**, 15–28.

Kiessling, L., Radbruch, J. and Schaube, S. (2018). The impact of self-selection on performance. *IZA Discussion Papers*, (11365).

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**, 409–426.

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

Kurzban, R., Duckworth, A., Kable, J. W. and Myers, J. (2013). An Opportunity Cost Model of Subjective Effort and Task Performance. *Behavioral and Brain Sciences*, **36** (6), 661–679.

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*, **87** (3), 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.

Pokorny, K. (2008). Pay–But Do Not Pay too Much: An Experimental Study on the Impact of Incentives. *Journal of Economic Behavior & Organization*, **66** (2), 251–264.

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.

Rosaz, J., Slonim, R. and Villeval, M. C. (2016). Quitting and Peer Effects at Work. *Labour Economics*, **39**, 55–67.

Waber, B., Magnolfi, J. and Lindsay., G. (2014). Workspaces that Move People. *Harvard Business Review*, **92** (10), 68–77.

Wagner, J. and Watch, D. (2017). *Innovation Spaces: The New Design of Work*. Brookings Institute.

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

A Appendix

A.1 Screenshot

Figure 3: Screenshot of Real-Effort Screen

A.2 Additional Figures and Tables

Table 6: Implicit (hourly) wage in Euro

	Piecerate-Low	Piecerate-High	Bonus-Easy	Bonus-Hard
Fix	1.73	8.91	6.67	0.00
Inet	1.26	8.05	6.41	0.32[a]
Free	1.62	8.56	6.95	0.00

Implicit wage is calculated by using the performance-dependent pay component (payoff without show-up fee) and scaling it up to an hourly wage. Surf time is working time, i.e., in Inet and Fix working time is fixed to 40 minutes. In Free, subjects can work less than 40 minutes.

[a] One subject achieved the target of 100. Without this subject, the implicit wage is 0.00.

Table 7: Regression of Output on Work Environments

	Over All Incentives	
	(1)	(2)
Inet	-7.97***	-8.32***
	(2.06)	(1.88)
Free	-15.45***	-15.17***
	(2.05)	(1.87)
Constant	58.79***	26.40***
	(1.46)	(5.28)
Controls	No	Yes
N	571	568
R^2	.091	.24
(Inet and Free = 0)	0.00	0.00
(Inet = Free)	0.00	0.00

Table presents least squares regression using output as dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust Standard errors in parentheses. For some observations controls are missing, due to some subjects who refused to answer some of the sociodemographic questions. Controls: ability, age, gender.

Table 8: Performance differences within Work-Environments

	Fix		Inet		Free	
	(1)	(2)	(3)	(4)	(5)	(6)
Piecerate-Low	-1.84	-6.11**	-11.82***	-12.97***	-16.14***	-15.80***
	(2.89)	(2.40)	(4.17)	(3.63)	(4.94)	(4.67)
Bonus-Easy	0.68	-3.97**	-0.50	-4.09	-6.41	-8.61**
	(2.72)	(2.00)	(3.27)	(2.97)	(4.21)	(4.17)
Bonus-Hard	-1.20	-4.48*	0.74	-1.76	-13.66***	-12.92***
	(3.14)	(2.53)	(3.89)	(3.57)	(4.78)	(4.58)
Constant	59.40***	23.86***	53.69***	24.65***	52.35***	25.95**
	(1.98)	(5.13)	(2.38)	(9.31)	(3.00)	(10.48)
Controls	No	Yes	No	Yes	No	Yes
N	188	188	190	189	193	191
R^2	.0047	.4	.067	.26	.068	.15
(PR-L vs Bonus-Easy)	0.37	0.35	0.01	0.01	0.05	0.14
(PR-L vs Bonus-Hard)	0.84	0.54	0.01	0.01	0.65	0.58
(Bonus-Easy vs Bonus-Hard)	0.54	0.84	0.74	0.51	0.13	0.38
Joint test of all vs. PR-H	0.81	0.06	0.02	0.00	0.00	0.00

Table presents least squares regression using Performance as dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust Standard errors in parentheses. For some observations controls are missing, due to subjects not answering the sociodemographic questions. Controls: ability, age, gender.

Table 9: Mann-Whitney test for comparisons between work environment

	Overall Fix	Piecerate-Low Fix	Piecerate-High Fix	Bonus-Easy Fix	Bonus-Hard Fix
Inet	0.0000	0.0010	0.0580	0.0068	0.3578
Free	0.0000	0.0001	0.1614	0.0001	0.0002
(Inet = Free)	0.0052	0.3422	0.9511	0.0372	0.0052

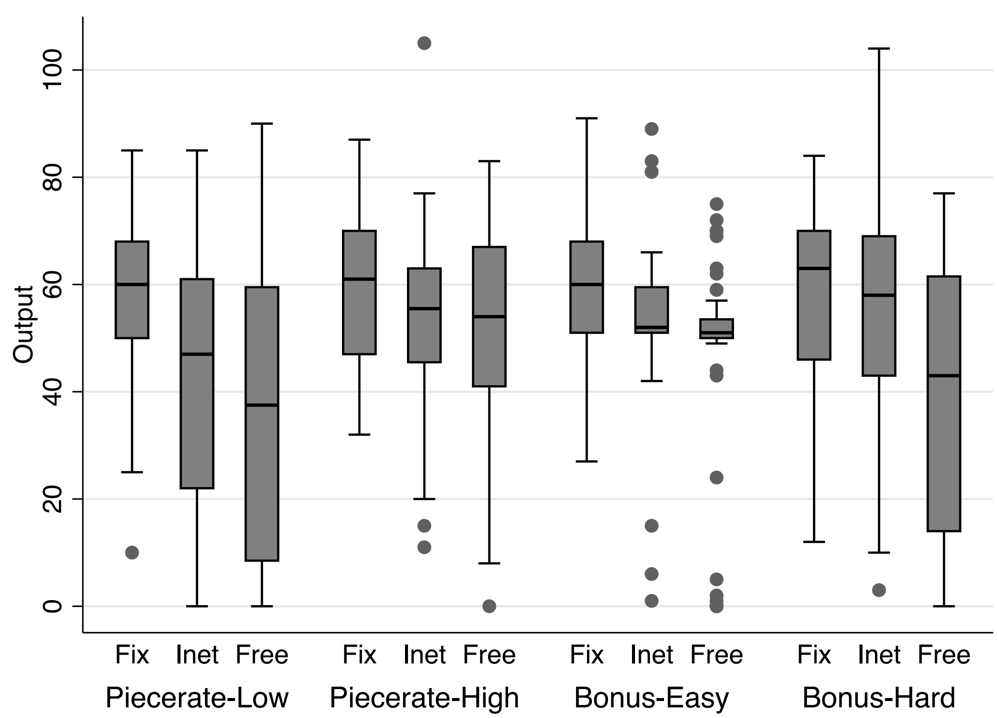
Table shows p-values of a Mann-Whitney u-test, which tests whether two independent samples were selected from populations with same distributions. Table is built analog to the tests of the regression coefficients in the corresponding regression tables.

Table 10: Mann-Whitney u-test for comparison of incentive schemes

	Fix	Inet	Free
	Piecerate-High		
Piecerate-Low	0.5866	0.0311	0.0050
Bonus-Easy	0.9191	0.5447	0.2055
Bonus-Hard	0.9644	0.3960	0.0176
(PR-L vs Bonus-Easy)	0.5231	0.0749	0.2594
(PR-L vs Bonus-Hard)	0.9644	0.3960	0.0176
(Bonus-Easy vs Bonus-Hard)	0.9571	0.1727	0.4343

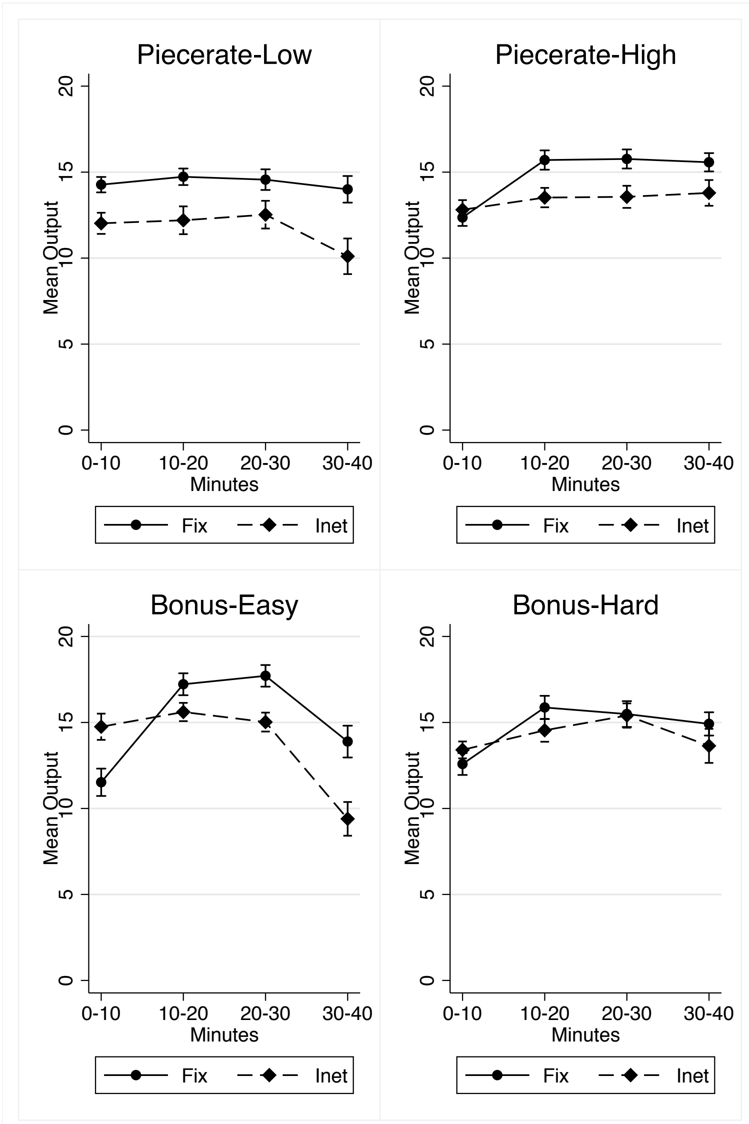
Table shows p-values of a Mann-Whitney u-test, which tests whether two independent samples were selected from populations with same distributions. Table is built analog to the tests of the regression coefficients in the corresponding regression tables.

Figure 4: Boxplot of Outputs in the different Treatments



Bold lines give the median outputs, boxes the 25th and 75th quartiles, and whiskers the 1.5xIQR. Circles present outliers, i.e., single observations outside of the whiskers.

Figure 5: Mean Output over Time



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

Online Appendix

June 23, 2018

Contents

A Examples of Experiments with Outside Options

Table 1 lists real-effort experiments which have treatments with and without outside options. Table 2 lists real-effort experiments with outside options. These papers do not manipulate the presence of the outside option.

Table 1: Economic Experiments with Manipulation of Outside Options

Author	Manipulate Outside Option	Outside Option(s)	Real-Effort	Task	Incentive Scheme	Research Topic
Corgnet *et al.* (2015d)	Yes	None, Internet surfing	Yes	Arithmetic summation task	Fixed wage + piecerate; fixed wage + team-incentive	Reaction to incentive schemes with internet as real-leisure option and without.
Dickinson (1999)	Yes	None, Stop working and Leave	Yes	Typing text	Various fixed wage + piecerate combinations	Individual wage elasticities with work intensity and leisure option.
Eckartz (2014)	Yes	None, Paid pause button	Yes	Letter Puzzle, arithmetic summation task	Fixed-wage; piecerate; tournament	Reaction to incentive schemes across task enjoyability and access to outside option.
Erkal *et al.* (2018)	Yes	Paid pause button, Stop working and leave, second effort task	Yes	encryption task	tournament	Reaction to tournament incentives with and without outside options.
Koch and Nafziger (2016)	Yes	None, Internet surfing	Yes	Counting zeros in tables	Principal-Agent: Agent receives fixed wage; Principal piecerate	Gift-exchange game when agents have access to internet or not.

Table 2: Economic experiments with outside option

Author	Manipulate Outside Option	Outside Option(s)	Real-Effort	Task	Incentive Scheme	Research Topic
Abeler *et al.* (2011)	No	Stop working and Leave	Yes	Counting zeros in tables	Lottery between fixed payment and piecerate	Compare quitting behavior across wages.
Berger *et al.* (2013)	No	Paid pause button	Yes	Counting sevens in tables	Bonus based on rating by supervisor	Effort of subjects working under performance appraisals based bonuses with and without forced distribution.
Blumkin *et al.* (2012)	No	With consumption goods paid pause button	Yes	Multiplication task	Consumption goods	Experimental test of the equivalence of wage and consumption taxes.
Charness *et al.* (2014)	No	Reading magazines	Yes	decoding one digit numbers to letter	Flat wage	Compare performance across treatments with and without feedback about relative performance (and possibilities of sabotage).
Corgnet *et al.* (2015f)	No	Internet surfing	Yes	Arithmetic summation task	Flat or chosen by principal	Test of the effect of influence activities of agents on performance.
Corgnet *et al.* (2015a)	No	Internet surfing	Yes	Arithmetic summation task	Piecerate split between principal and agent	Interplay of goal-setting and monetary incentives.
Corgnet *et al.* (2015b)	No	Internet surfing	Yes	Arithmetic summation task	Flat wage, piecerate	Effect of firing threats on performance.

Table 2: Experiments with Outside Options

Author	Manipulate Outside Option	Outside Option(s)	Real-effort	Task	Incentive Scheme	Research Topic
Corgnet *et al.* (2015e)	No	Internet surfing	Yes	Arithmetic summation task	Principal Agent contract, share of agents production (with and without noise)	Compare production in noisy environment to production in environment without noise
Corgnet *et al.* (2015c)	No	Internet surfing	Yes	Arithmetic summation task	Team incentives or piecerate	Comparison of team incentives (with monitoring) and individual incentives
Eriksson *et al.* (2009)	No	Reading Magazines	Yes	Arithmetic summation task	Piecerate; tournament	Impact of incentives and relative performance feedback on performance.
(Falk and Huffman, 2007)	No	Stop working and Leave	Yes	Counting zeros in tables		Subjects had to fulfill work requirement, but could leave as soon as they were done.
Hayashi *et al.* (2013)	No	Preselected YouTube videos	Yes	Alphabetizing words	Flat payment if leisure option is chosen or piecerate	Study the reaction to different tax regimes.
Hammermann and Mohnen (2014)	No	Reading Magazines	Yes	Solving mathematical equations	Tournament with (non-) monetary prize	Compare subjects' performance in tournaments with monetary and non-monetary prizes

Table 2: Experiments with Outside Options

Author	Manipulate Outside Option	Outside Option(s)	Real-effort	Task	Incentive Scheme	Research Topic
Kajackaite (2015)	No	Paid pause button	Yes	Decoding letters	Piecerate + possibility of piecerate for NRA	Compare subjects performance in treatments with private piecerate and additional piecerate for NRA (possibly unknown, subjects can stay ignorant)
Kessler and Norton (2016)	No	Internet surfing	Yes	Typing strings	Piecerate	Performance of subjects, with wage decreases due to tax or wage cuts.
Mohnen *et al.* (2008)	No	Paid pause button	Yes	Counting sevens in tables	Team-incentives	Effect of information about the performance of the other team member on performance.
Rosaz *et al.* (2016)	No	Stop working and Leave	Yes	Arithmetic math task	Fixed wage + piecerate	Compare quitting behavior across treatments, varying presence of a peer.

References

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

Berger, J., Harbring, C. and Sliwka, D. (2013). Performance Appraisals and the Impact of Forced Distribution – An Experimental Investigation. *Management Science*, **59** (1), 54–68.

Blumkin, T., Ruffle, B. J. and Ganun, Y. (2012). Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment with Real Goods. *European Economic Review*, **56** (6), 1200–1219.

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

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 Rassenti, S. (2015b). Firing Threats: Incentive Effects and Impression Management. *Games and Economic Behavior*, **91**, 97–113.

—, Hernan-Gonzalez, R. and Rassenti, S. (2015c). Peer Pressure and Moral Hazard in Teams: Experimental Evidence.

—, Hernán-González, R. and Schniter, E. (2015d). Why Real Leisure Really Matters: Incentive Effects on Real Effort in the Laboratory. *Experimental Economics*, **18** (2), 284–301.

—, Hernan-Gonzalez, R. *et al.* (2015e). Revisiting the Tradeoff between Risk and Incentives: The Shocking Effect of Random Shocks.

—, Martin, L., Ndodjang, P. and Sutan, A. (2015f). On the Merit of Equal Pay: When Influence Activities Interact with Incentive Setting.

DellaVigna, S. and Pope, D. (2018). What Motivates Effort? Evidence and Expert Forecasts. *The Review of Economic Studies,*, **85** (2), 1029–1069.

Dickinson, D. L. (1999). An Experimental Examination of Labor Supply and Work Intensities. *Journal of Labor Economics*, **17** (4), 638–670.

Eckartz, K. (2014). Task Enjoyment and Opportunity Costs in the Lab: The Effect of Financial Incentives on Performance in Real Effort Tasks. *Jena Economic Research Papers*.

ERIKSSON, T., POULSEN, A. and VILLEVAL, M. C. (2009). Feedback and Incentives: Experimental Evidence. *Labour Economics*, **16** (6), 679–688.

ERKAL, N., GANGADHARAN, L. and KOH, B. H. (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review*, **101**, 528 – 545.

FALK, A. and HUFFMAN, D. (2007). Studying Labor Market Institutions in the Lab: Minimum Wages, Employment Protection, and Workfare. *Journal of Institutional and Theoretical Economics*, **163** (1), 30–45.

HAMMERMANN, A. and MOHNEN, A. (2014). The pric(z)e of hard work: Different incentive effects of non-monetary and monetary prizes. *Journal of Economic Psychology*, **43**, 1–15.

HAYASHI, A. T., NAKAMURA, B. K. and GAMAGE, D. (2013). Experimental Evidence of Tax Salience and the Labor–Leisure Decision. *Public Finance Review*, **41** (2), 203–226.

KAJACKAITE, A. (2015). If I Close my Eyes, Nobody Will Get Hurt: The Effect of Ignorance on Performance in a Real-Effort Experiment. *Journal of Economic Behavior & Organization*, **116**, 518–524.

KESSLER, J. B. and NORTON, M. I. (2016). Tax Aversion in Labor Supply. *Journal of Economic Behavior and Organization*, **124**, 15–28.

KOCH, A. K. and NAFZIGER, J. (2016). Gift Exchange, Control, and Cyberloafing: A Real-Effort Experiment. *Journal of Economic Behavior & Organization*, **131**, 409–426.

KURZBAN, R., DUCKWORTH, A., KABLE, J. W. and MYERS, J. (2013). An Opportunity Cost Model of Subjective Effort and Task Performance. *Behavioral and Brain Sciences*, **36** (6), 661–679.

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.

ROSAZ, J., SLONIM, R. and VILLEVAL, M. C. (2016). Quitting and Peer Effects at Work. *Labour Economics*, **39**, 55–67.

B 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. I is an indicator function which is $I = 1$ if the agent is intrinsically motivated and $I = 0$ if not.[1]

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.[2,3] 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 interval $[0, E]$ and for simplicity $c(0, i) = 0\ \ \forall i \in \{i_{Fix}, i_{Inet}, i_{Free}\}$. 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., $\lim_{e \to E} \frac{\partial c(e,i)}{\partial e} = \infty$.[4]

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

[2] This includes, for example, versions of $c(e,i)$ like in Koch and Nafziger (2016), where $c(e,i) = i \times \tilde{c}(e)$ and the parameter i differs between work environments, i.e., effort costs increase when alternative actions are present. Therefore $i_{Fix} \leq i_{Inet} \leq i_{Free}$. 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.* (2015d).

[3] Implicit effort costs in this setup represent the forgone utility of not allocating the effort or time to other activities.

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

Piecerate Incentives:
We first discuss the two piecerate incentive schemes. Under those incentive schemes, the maximization problem leads to the following first-order condition, stating that agents supply effort as long as the marginal benefit of effort is higher than the marginal cost of effort:

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

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, our work environment manipulation increases 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. Therefore, the optimal effort level e^* and its associated output decrease.

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 piecerate. However, it could be 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 marginal extrinsic incentives only 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 $\hat{e}$ be the effort level that is needed to meet the target, i.e., $g = f(\hat{e})$. 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 losing any benefit. Similarly, it is easy to see that no effort above $\hat{e}$ can be optimal. The agent considers either exerting exactly the effort level $\hat{e}$, which is needed to reach the target ($g = f(\hat{e})$), or npt exerting any 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.

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

This 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 a subject needs in order to reach the target g.

This implies that for Bonus-Easy we would expect only few outputs above 50, since any additional output above the target would only be driven by workers who 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 who exert a positive effort level, which leads to an output below the target. Especially in 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 *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 who exert an effort level $\hat{e}$. Some of these 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, who already exerted an effort level below $\hat{e}$ will also decrease their effort.

C Structural Estimation

C.1 Parametrization of Conceptual Framework

We have to parametrize the functions in order to estimate the model structurally. We will focus on the piecerate treatments, as those have inner solutions and can be easily estimated. In general, our parametrization closely follows common specifications in the real-effort literature (e.g., DellaVigna and Pope, 2018). The production technology $y = f(e) = e$ translates effort to output. The fixed wage, i.e., a lump-sum payment, is represented by $\bar{w}$. The intrinsic motivation is represented by $I\delta(y)$. We parametrize this in a linear way, i.e., all agents are intrinsically motivated by $\delta(y) = s * e$.

The effort costs are represented by $C(e, i)$, which includes the explicit as well as implicit effort costs. We assume that the parameter i increases the marginal effort costs depending on the outside options available to the agent. Setting up the effort costs in this setup is crucial, since our treatment variations are changing the implicit effort costs. There are two possible ways how effort costs could change. First, physical effort costs could potentially be multiplied, i.e., every effort unit is more expensive as an agent faces opportunity costs. This would multiply the physical effort costs, for example, with a factor o_i. Second, agents face an additional cost of effort for every unit, i.e., the foregone utility of spending this effort differently. Therefore, we add another term to the cost function, for simplicity a linear term $a_i * e$. These two possibilities lead to the following effort costs $C(e, i)$, where our treatments might potentially change only o or a (see discussion below):

$$C(e, i) = exp(o_i)c(e) + a_i e$$

.

For $c(e)$, we can use two versions of effort costs commonly used in the literature: a power cost function and an exponential cost function.

1. Power Cost Function:

$$c(e) = \exp(k)\ \frac{e^{1+\gamma}}{1+\gamma}$$

The power cost function has a constant elasticity with respect to the value of effort (i.e., $s + p$) of $1/\gamma$ and can be scaled with some (positive) parameter $exp(k)$.[5]

[5] $exp(k)$ is used to ensure that the parameter is positive.

2. Exponential Cost Function:
A natural alternative is a function with a decreasing elasticity, one function with such a structure being the exponential cost function:

$$c(e) = \exp{(k)}\ \frac{\exp{(\gamma e)}}{\gamma}.$$

Given this parametrization, the agent solves the following maximization problem:

$$\max_{e>0} u(e) = \bar{w} + (p+s) * e - exp(o_i)c(e) - a_i * e$$

This will lead to a first-order condition which holds with equality due to the properties of $c(e)$. It is setting the marginal costs of effort equal to the marginal benefit. Given the two parametrizations of $c(e)$, this leads to the following solutions for the optimal effort. For **power costs**:

$$log(e^*) = \frac{1}{\gamma}[log(p+s-a_i) - k - o_i] \tag{2}$$

and for the **exponential cost function**:

$$e^* = \frac{1}{\gamma}[log(p+s-a_i) - k - o_i] \tag{3}$$

C.2 Structural estimation

To estimate the above model structurally with non-linear least squares, we need to add some noise term. If we add a noise term to the cost of effort function, the cost function $C(e,i)$ of worker j is

$$C_j(e,i) = exp(o_i)c(e) * exp(-\gamma * \epsilon_j) + a_i e$$

.

This will lead to the following two equations:

$$log(e_j) = \frac{1}{\gamma}[log(p+s-a_i) - k - o_i] + \epsilon_j \tag{4}$$

$$e_j = \frac{1}{\gamma}[log(p+s-a_i) - k - o_i] + \epsilon_j. \tag{5}$$

For the case without opportunity costs, $a_i = 0$ and $o_i = 1$, and both first-order conditions have three unknown parameters (γ, s, k). In each work environment with opportunity costs, both first order conditions have two additional unknown parameters (a_{Inet}, o_{Inet} and a_{Free}, o_{Free}); the work environments with opportunity costs add four additional parameters in total.

In order for our model to be identified, we use the following restrictions: First, we only allow changes of opportunity costs to matter due to changes in a_i, i.e., each effort unit produces some additional costs. This idea is also in line with the standard idea of how opportunity costs should enter. This shifts the marginal effort costs upwards and, by this, potentially decreases effort. We therefore restrict $o_i = 1$ in all environments.[6, 7] Furthermore, we can add an additional parameter to account for potential heterogeneous effort costs due to differentiability. For this, we can multiply the effort costs with an additional parameter $exp(ability) * performance\_trial$. This allows for heterogeneous effort cost functions due to ability.

C.3 Results

Table 3 presents the results of the estimation. We first focus on the main parameters of interest, a_{Inet} and a_{Free}. We observe that both parameters are positive across all specifications. In both environments, subjects have to pay an additional cost for each unit produced. However, in our main specifications (1) and (4), both parameters are not significant. Only if we allow ability to enter in (3) and (6) parameters become significant. This is due to the high variance in our data. Overall, the coefficients are in line with the hypothesis that opportunity costs shift the effort costs upwards and that the shift is slightly higher in FREE than in INET. Our estimates also show that implicit costs are of a similar size as the intrinsic motivation parameter. The effect of introducing opportunity costs therefore has a similar effect in terms of magnitude as setting off intrinsic motivation.

To check the fit of our model, we tabulate the predicted output of the model and the actual output in Table 4, using the model in column (1). In Figure 1, we plot the results of column (1) to illustrate the results and the mechanics. In INET and FREE, the marginal

[6] As a robustness check, we allow one additional parameter that is the same for both opportunity cost environments.

[7] We estimate a version where we allow only for changes in a_i. We observe a very low elasticity with respect to changes in p in FIX. A model which tries to fit these moments will therefore estimate a very high γ. Multiplying this kind of effort costs curve with a parameter smaller than 1 will reduce effort (as marginal costs are higher), yet will be unable to match the elastic response to effort in the other two environments.

Table 3: STRUCTURAL PARAMETERS OF EFFORT COSTS

	Power Costs			Exponential Cost		
	(1)	(2)	(3)	(4)	(5)	(6)
γ	6.1402	3.7928	3.5236	0.1659	0.1925	0.0835
	(8.2545)	(5.6732)	(2.5826)	(0.1854)	(0.2733)	(0.0571)
s	14.4257	36.7609	8.5172*	16.8910	16.8010	11.3334**
	(20.9882)	(150.3160)	(5.0409)	(17.6572)	(21.9920)	(4.6492)
a_{Inet}	15.6364	36.6680	8.4170	17.4483	17.8837	9.1744**
	(22.9803)	(148.1987)	(5.4080)	(20.2475)	(24.2706)	(3.9922)
a_{Free}	16.3858	38.4273	10.1634**	18.3353	18.4262	10.9095***
	(21.2072)	(149.5680)	(5.0685)	(19.1204)	(23.2552)	(3.8420)
k	-21.7657	-11.5388	-8.7176	-6.5898	-8.1466	0.5525
	(32.3766)	(24.5867)	(8.3318)	(10.1751)	(15.2176)	(1.7276)
o		-1.0136			0.0976	
		(3.7866)			(0.5420)	
ability			-0.5118			-0.4680
			(0.3772)			(0.3190)
N	283	283	281	287	287	285
R^2	.18	.18	.26	.15	.15	.33

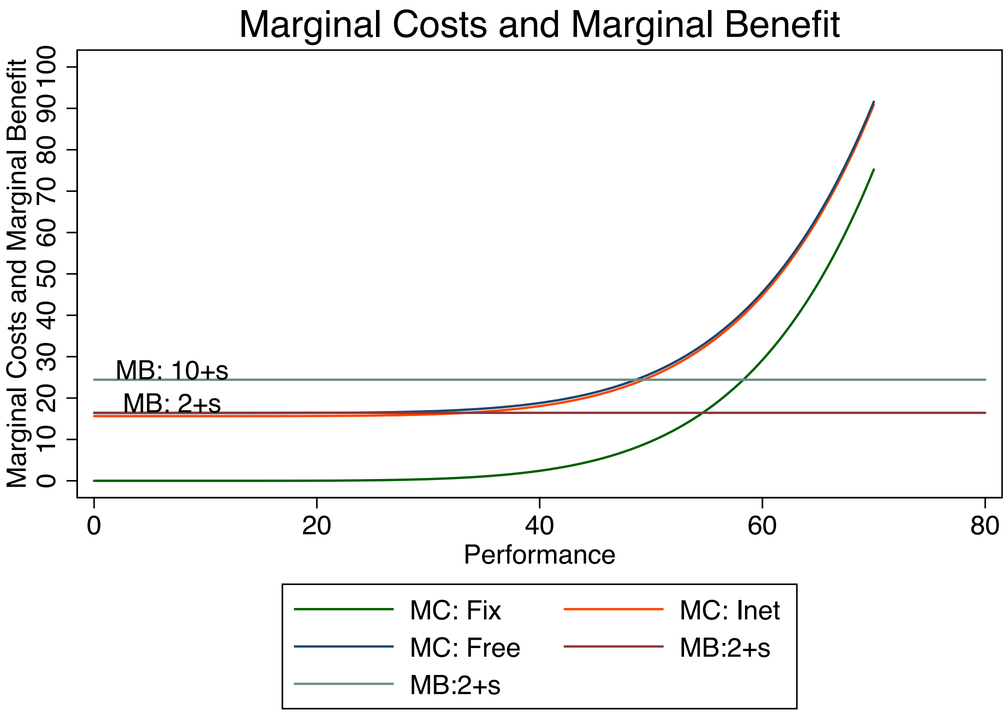
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Table presents structural estimates of 4 and 5 using non-linear least squares. Standard errors in parentheses.

cost curve is shifted upwards due to a_{Inet} and a_{Free}. As the marginal benefits are constant across the environments, this reduces both the observed and estimated outputs.

Table 4: Output and predicted output

	FIX		INET		FREE	
	PIECERATE-LOW	PIECERATE-HIGH	PIECERATE-LOW	PIECERATE-HIGH	PIECERATE-LOW	PIECERATE-HIGH
	Mean	Mean	Mean	Mean	Mean	Mean
Predicted	54.63255	58.27958	33.32347	49.34271	20.49268	48.63174
Output	57.5625	59.40426	41.87234	53.6875	36.20833	52.34694
Observations	48	47	47	48	48	49

Figure 1: Marginal Costs and Marginal Benefit with Power Costs



Bonus-based incentives

We can use our estimated parameters and calculate a prediction for the bonus-based treatments using some simplifications and assumptions. We use the estimates of column(1) and show the results of this exercise in Table 5. For Fix, we can simply drop the marginal piece rate incentives. In the case Bonus-easy, Case 2a from above is fulfilled. The intrinsic motivation equilibrium predicts effort slightly above 50. For Bonus-hard, Case 2b from above is fulfilled. However, reaching the target is too costly. Therefore we get the same prediction as before.

For both environments Inet andFree intrinsic motivation is not high enough to induce an equilibrium with pure intrinsic motivation, as $s < a$. Therefore we only have to check whether the subjects target the goal or not. We can simply compare, costs and benefits of working until the goal is reached. We therefore predict an output of 50 for Bonus-easy and 0 for Bonus-hard.

For the fix environment, our model makes predictions that are generally in line with our results. For the other two environments, the point predictions are also close to the observed outcome in Bonus-Easy. The model, however, fails to predict the observed outcome in the Bonus-Hard environments. Generally, the model makes very sharp predictions, although it does not take into account the heterogeneity in intrinsic motivation.

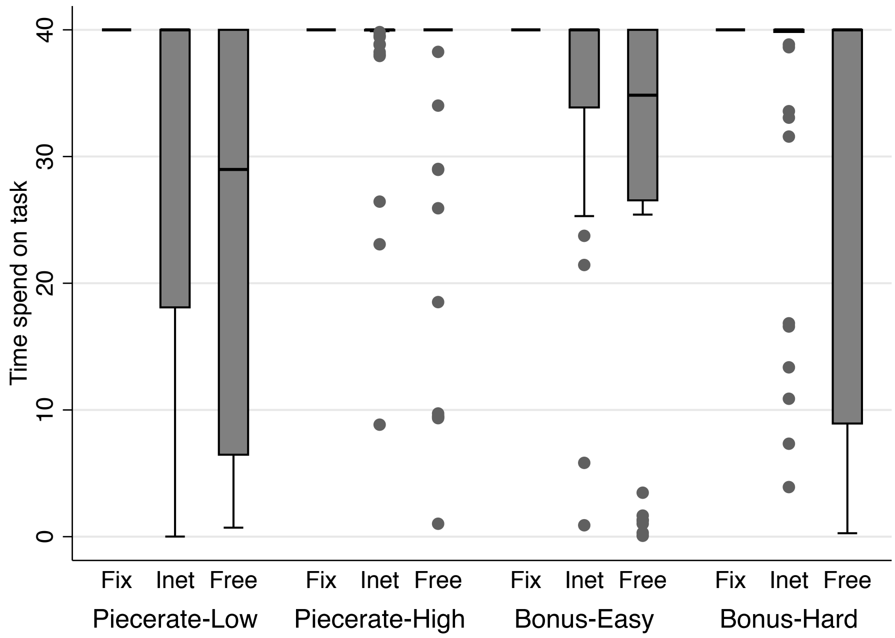
Table 5: Output and predicted output

	Fix		Inet		Free	
	Bonus-Easy	Bons-Hard	Bonus-Easy	Bons-Hard	Bonus-Easy	Bons-Hard
	Mean	Mean	Mean	Mean	Mean	Mean
Predicted	53.48947	53.48947	50	0	50	0
Output	60.08889	58.20833	53.1875	54.42553	45.9375	38.6875
Observations	45	48	48	47	48	48

D 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 an outside option. In the following, we show that subjects actually used the outside option. In addition, we show that our treatment differences mainly result from the time worked on the task (extensive margin) rather than the speed (intensive margin).

Figure 2: Boxplot of time spent working on the task



Bold lines give the median outputs, boxes the 25th and 75th quartiles, and whiskers the 1.5xIQR. Circles present outliers, i.e., single observations outside of the whiskers.

Across all treatments, the time spent working on the task significantly correlated with the output ($\rho = 0.6048, p < 0.01$). Figure 2 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 $p < 0.01$, Wilcoxon signed-rank test). Table 6 presents OLS regressions which show how our work environments affect the time subjects work on the task, analogously to the treatment effect tables in the paper.

In a next step we adjust the output by the total time an individual worked on the task. If the time spent working on the task is the main driver of the treatment differences, we should not observe significant differences between treatments anymore. Figure 3 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 $p < 0.01$ to $p = 0.1277$; in Bonus-Easy

Table 6: Treatment effects for Time working on the task

	Over All Incentives		Piecerate-Low		Piecerate-High		Bonus-Easy		Bonus-Hard	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Inet	-4.88***	-5.23***	-9.36***	-8.22***	-1.48*	-1.73*	-4.59***	-5.00***	-4.16***	-4.20***
	(0.74)	(0.75)	(2.03)	(1.89)	(0.77)	(0.96)	(1.22)	(1.28)	(1.42)	(1.47)
Free	-10.82***	-10.64***	-15.79***	-14.06***	-4.62***	-4.91***	-10.07***	-9.66***	-12.94***	-12.73***
	(1.05)	(1.05)	(2.35)	(2.42)	(1.48)	(1.49)	(1.88)	(1.85)	(2.30)	(2.38)
Constant	40.00***	35.10***	40.00***	18.08**	40.00	39.36***	40.00***	46.62***	40.00***	39.26***
	(0.00)	(3.12)	(0.00)	(7.73)	(.)	(3.85)	(0.00)	(7.65)	(0.00)	(7.16)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	571	568	143	142	144	142	141	141	143	143
R^2	.16	.18	.22	.29	.076	.084	.17	.22	.2	.21
(Inet and Free = 0)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(Inet = Free)	0.00	0.00	0.04	0.06	0.06	0.07	0.02	0.03	0.00	0.00

Table presents results from an OLS estimation with output as dependent variable. * p < 0.1, ** p < 0.05, *** p < 0.01. Robust Standard errors in parentheses.

from $p < 0.01$ to $p = 0.8592$, in Bonus-Hard from $p < 0.01$ to $p = 0.2579$; and in Piecerate-High from $p = 0.1636$ to $p = 0.3316$ (all Kruskal-Wallis test).[8]

We additionally analyze the time subjects need to complete one screen to measure effort adjustments at the intensive margin. Table 7 gives the average and median time needed per screen. Testing the median time per screen reveals no significant differences in the medians for all incentive schemes, except for Piecerate-High (all $p > 0.446$, median test). For Piecerate-High, the median does differ across work environments ($p = 0.098$, median test).[9] Comparing the means gives no significant differences for any incentive scheme (all $p > 0.2250$, Kruskal-Wallis test).

[8]The Kruskal-Wallis test is a multi-sample generalization of the two-sample Mann-Whitney u-test.

[9]One possibility which influences the time subjects need to complete a screen is the existence of opportunity costs and the salience of those in a given situation. See Kurzban *et al.* (2013) for a discussion in Pyschology.

Figure 3: Output, adjusted by working time

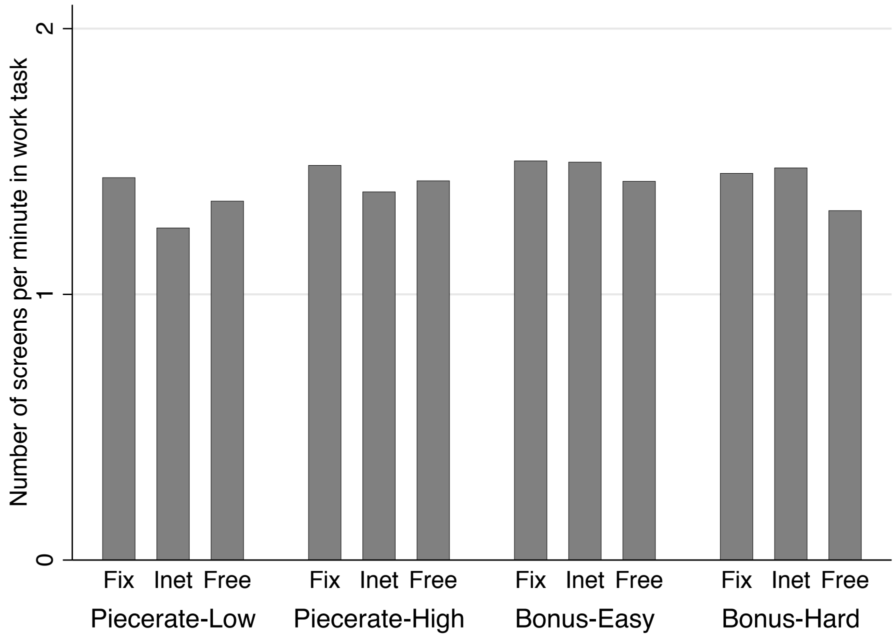


Table 7: Time per Screen

		Over All Incentives	By Incentive Scheme Piecerate-Low	Piecerate-High	Bonus-Easy	Bonus-Hard
Fix	Mean	45.14	47.60	42.88	42.17	47.67
	Median	39.34	40.00	39.34	40.00	38.10
	SD	22.36	30.95	11.49	11.79	27.30
	N	188	48	47	45	48
Inet	Mean	46.48	52.89	47.65	42.04	43.70
	Median	40.96	41.90	41.39	38.83	41.74
	SD	18.51	25.20	20.79	10.05	13.08
	N	188	45	48	48	47
Free	Mean	45.69	49.20	44.24	40.62	48.49
	Median	40.96	41.90	41.39	38.83	41.74
	SD	18.21	25.03	12.98	11.67	19.01
	N	185	47	48	44	46

SD: standard deviation, N: number of independent observations

E Explaining the usage of the outside option

We repeat the analysis from the paper and check whether our questionnaire measures can explain the usage of the outside option. Table 8 presents results for the intensive margin, i.e., the time subjects work on the task, and Table 9 presents the effects on the extensive margin, i.e., on the probability of using the outside option. For the latter we present marginal effects using a logit model with an indicator if someone used the outside option. Overall, results are in line with the analysis in the paper for both the time subjects spent working on the task and the probability of using the outside option.

Table 8: Time spent working

	(1)	(2)	(3)
Inet x CRT score	0.0753		0.5809
	(0.6505)		(0.6573)
Free x CRT score	-1.8688**		-1.7050*
	(0.8899)		(1.0040)
Inet x Conscientiousness		1.0099**	0.9303**
		(0.3992)	(0.4246)
Free x Conscientiousness		0.6986	0.2125
		(0.5936)	(0.6767)
Inet x Ability			0.5703
			(0.4986)
Free x Ability			0.3815
			(0.6483)
Treatment FE	Yes	Yes	Yes
Gender and Age	No	No	Yes
Big 5 (w/o Cons), Risk	No	No	Yes
N	383	383	380
R^2	.14	.14	.2

Table presents least squares regression using the time worked on the task as dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust Standard errors in parentheses. Big 5 (w/o Cons.) controls for the other Big 5 traits and risk for general risk attitudes.

Table 9: PROBABILITY OF USING OUTSIDE OPTION

	(1)	(2)	(3)
used outside option			
INET x CRT score	0.1343		0.1027
	(0.1309)		(0.1427)
FREE x CRT score	0.3374**		0.2100
	(0.1421)		(0.1651)
INET x Conscientiousness		-0.1764**	-0.1271
		(0.0872)	(0.0899)
FREE x Conscientiousness		-0.2006**	-0.0936
		(0.0970)	(0.1183)
INET x Ability			-0.1110
			(0.1078)
FREE x Ability			0.0600
			(0.1007)
Constant	-0.7592**	0.7431	2.9188
	(0.3825)	(0.5236)	(1.8338)
Treatment FE	Yes	Yes	Yes
Gender and Age	No	No	Yes
Big 5 (w/o Cons), Risk	No	No	Yes
N	383	383	380

Table presents odd ratios from a logit model with an indicator if someone used the outside option as dependent variable. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust Standard errors in parentheses. Big 5 (w/o Cons.) controls for the other Big 5 traits and risk for general risk attitudes.

F Experimental Instructions

F.1 General Information

General Instructions

Thank you for participating in today's study. Please read the following instructions carefully. If you have questions, you can can ask them at the end of the introduction. To carry out the study, it is very important that you do not communicate with other participants. Therefore, you are not allowed to talk to others. If you communicate with another participant regardless of this, you will have to leave the experiment and will receive no payment.

In this study, you will have the possibility to earn money. The payment at the end of the study is done individually and no other participant will know how much you earned in this study.

Instructions for the task

The task in this study is to set as many sliders as possible to the middle position (position 50) in a given time. Each slider is located at the left end (position 0) and can be moved in steps of 1 to the right end of the scale (position 100). The current position of the slider is displayed to the right of the scale. Please use your mouse to move the slider on the scale as desired. Only when all sliders on the screen are located at the center (position 50) will a red button appear. By pressing this button, you confirm that all sliders are in the middle and you will earn a point.

Please note: You will only earn a point if all sliders are in the middle and you pressed the red button. The task will start simultaneously for all participants. You can see your personal score at the top right corner of the screen. We will now start with a trial round. In this trial round, you can familiarize yourself with the task. Following the trial round, you will receive further information.

F.2 Treatment-Specific Instructions

Performance-oriented remuneration

We ask you to do your job carefully. Please try to finish as many screens with sliders as possible within the next 40 minutes. It is not possible to terminate the task before that time is up, since the payment will only be done at the end of the experiment.[10] Generally, the more points you earn, the higher the payment, which you will receive from us immediately afterwards in cash. The following will be applied:

[**PR Treatments:**]

- You will receive a basic wage of 10 Euros. That means that you will earn at least 10 euros for the 40 minutes.
- In addition to your basic wage, you will receive a bonus payment. The size of this bonus payment is based on the number of points you collect:

  For each point you will get an additional 2 [10] cents.

[**Bonus Treatments:**]

- You will receive a basic wage of 10 Euros. That means that you will earn at least 10 euros for the 40 minutes.
- In addition to your basic wage, you may receive a bonus payment of 5 [10] Euro. It depends on the number of points you accumulate whether you receive this bonus or not. We set a personal goal for you that is 50 points. If you do not reach this goal within 40 minutes, you will not receive the bonus. If you reach the goal, you will receive the bonus payment of 5 [10] Euros.
  - *Example:* You will receive a bonus of 5 [10] Euros as soon as you have collected 50 points or more (also, if you have collected, for example, 105 points). If you have accumulated less than 50 points, you will not receive the bonus.

You will only earn a point if all sliders are in the middle and you have pressed the red button.

[**Inet Treatments:**] During the next 40 minutes, you can also surf the Internet. You can access the internet by clicking the "Internet" button. If you click on this button, Internet Explorer will open. As long as you are on the Internet, your work will be interrupted. To continue working on the task, close Internet Explorer and click "Proceed". You can also interrupt your work several times and return to the task at any time. After 40 minutes, Internet Explorer closes automatically and you can no longer return to the task either.

[10] For Free treatments, this sentence was replaced by the following: You can stop working on the task at any time. If you decide to stop working on the task, you can collect your payment and the study is finished for you.

[**Free treatments:**] You can stop working on the task at any time. If you decide to stop working on the task, your payment is based on all the points you have earned up to this point.

G Screenshot and Implementation

Original screen resolution was 1920 x 1200 and is adjusted for the screenshots to 1024 x 768.

Figure 4: Screenshot of Real-Effort Screen in Fix

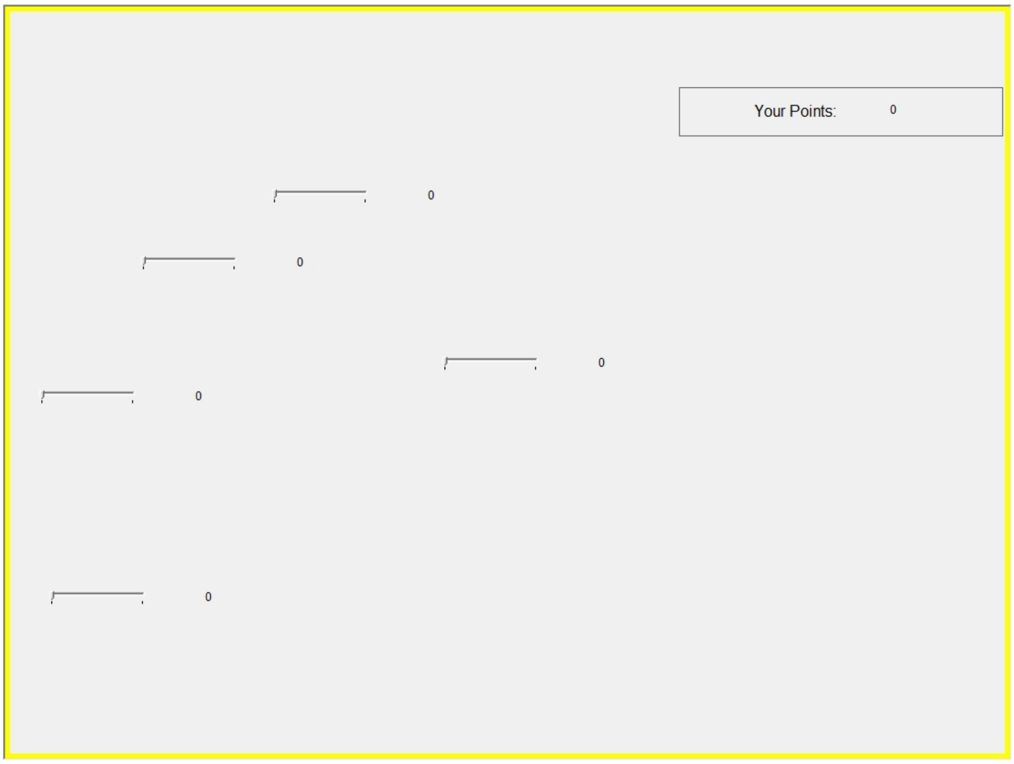


Figure 5: Screenshot of Real-Effort Screen in INET

Figure 6: Screenshot of Internet Access Screen in INET

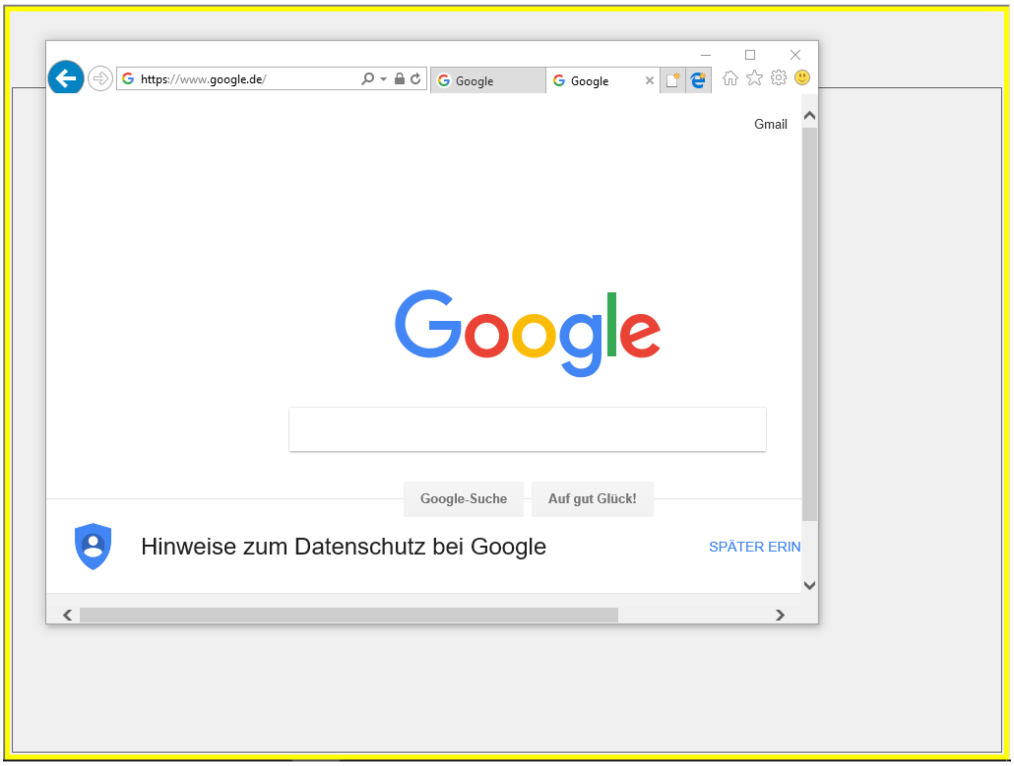


Figure 7: Screenshot of blocked Real-Effort Screen in INET

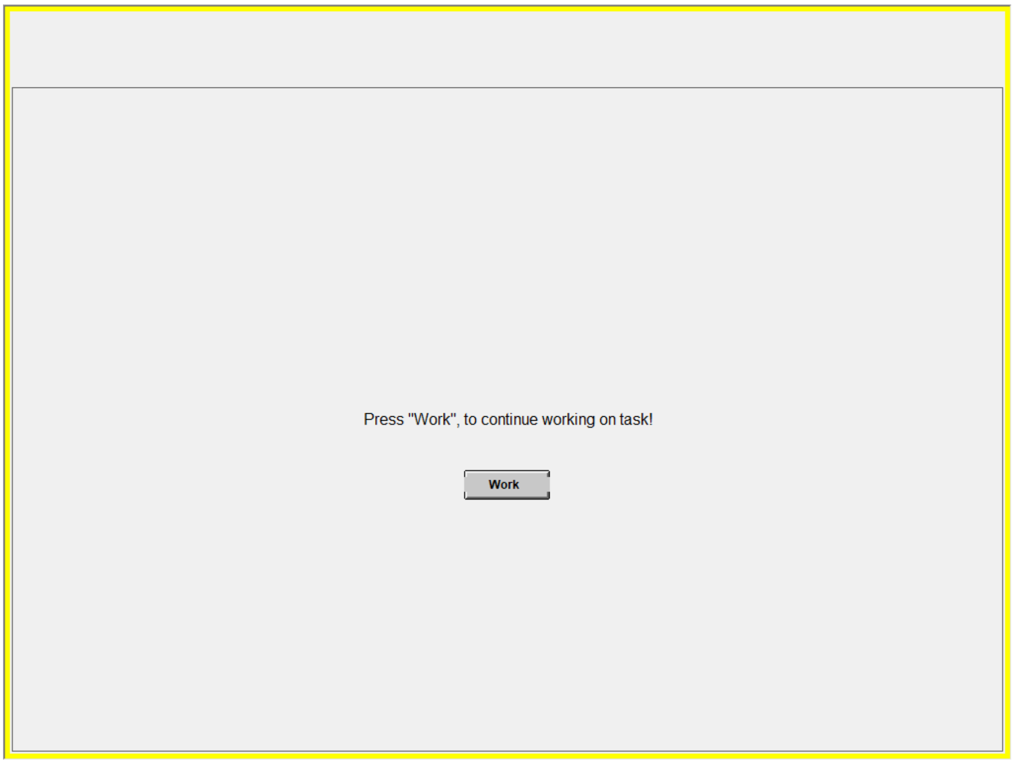
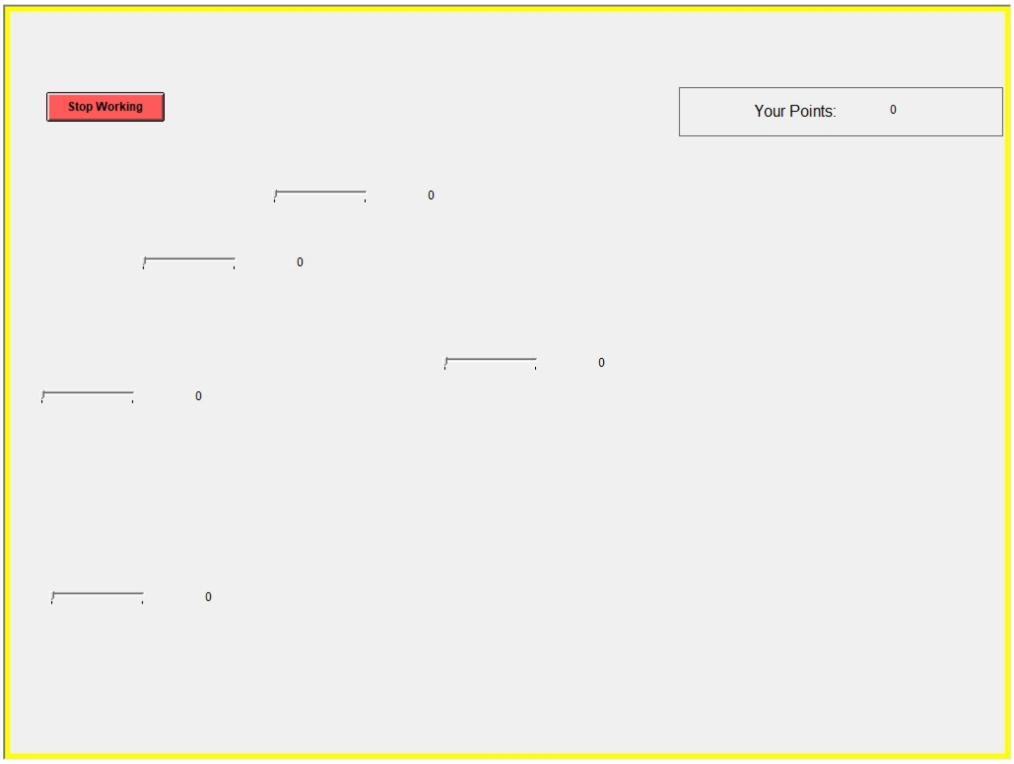


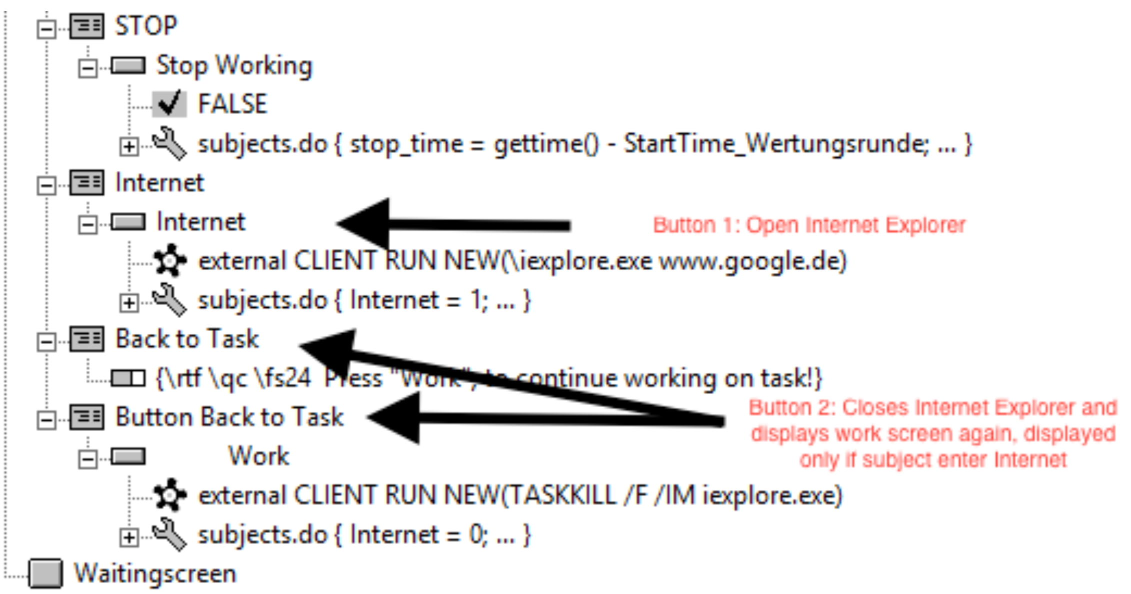
Figure 8: Screenshot of Real-Effort Screen in FREE



G.1 Implementation of INET and FREE

We implemented the INET environment by adding a button to the screen which allowed subjects to open Internet Explorer on their computer (see Figure 5 and Figure 9). This button calls an external program, i.e., in our case Internet Explorer. Furthermore, the button indicates that the subjects is online and switches an indicator variable *Internet* to 1. If the indicator variable is switched to 1 a box is displayed, which covers the real-effort task (see Figure 7). Thus, upon pressing the button, an Internet Explorer window opens, the slider task is blocked, and subjects can surf the web. If subjects want to return to the task, they can close the Internet Explorer or click on the zleaf window. Subjects would automatically return to zleaf, as it is still running in full screen mode. The blocking screen entailed a button that switched the indicator variable back to 0 and the real-effort task would be displayed again. The key "Alt" on the subjects' keyboard was disabled; therefore they could not switch between windows or close any other window except for the Internet Explorer window. Figure 9 gives details on the implementation. We implemented FREE in a similar way. The only difference was that, upon clicking the button 'Stop Working', subjects would leave the work stage instead of accessing the internet.

Figure 9: Implementation of INET in z-Tree



G.2 Payment and Procedures in Free

Subjects in Free could arrive during a given time window on a given day. Subjects entered the laboratory through the entrance and were quietly directed to the second room by an experimenter (see Figure 10). Registration took place in the second room and subjects would proceed with the experiment in their computer cabin in the first room. Cabins are separated by walls and subjects work behind a closed curtain (see Figure 11). All cabins are accessible without anyone being disturbed. Upon completion subjects are told to go back to the second room quietly and payment was done there in private.

Figure 10: Sketches of BonnEconLab

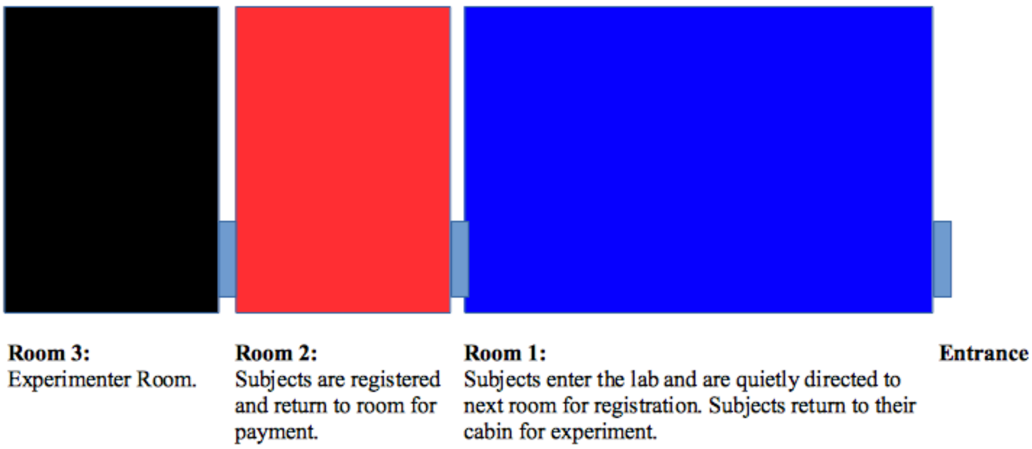


Figure 11: Photos of lab and sketches of laboratory room

(a) Computer cabins 1

(b) Computer cabins 2

(c) Computer cabins 3