Discussion Paper No. 358

Reference Points and Effort Provision

Johannes Abeler*
Armin Falk**
Lorenz Götte***
David Huffman****

*University of Nottingham
** University of Bonn
*** University of Lausanne
**** Swarthmore College

May 2011

Reference Points and Effort Provision[†]

Johannes Abeler

University of Nottingham

Armin Falk

University of Bonn

Lorenz Götte

University of Lausanne

David Huffman

Swarthmore College

November 7, 2009

A key open question for theories of reference-dependent preferences is what determines the reference point. One candidate is expectations: what people expect could affect how they feel about what actually occurs. In a real-effort experiment, we manipulate the rational expectations of subjects and check whether this manipulation influences their effort provision. We find that effort provision is significantly different between treatments in the way predicted by models of expectation-based reference-dependent preferences: if expectations are high, subjects work longer and earn more money than if expectations are low.

JEL classification: C91, D01, D84, J22.

Keywords: Reference Points, Expectations, Loss Aversion, Disappointment, Experiment

1 Introduction

Imagine two identical workers. One expected a salary increase of 10 percent but receives an increase of only 5 percent. The other receives the same 5 percent wage increase but had not expected an increase. The change in income is the same for both workers, but the first worker presumably feels less satisfied. Intuitively, many people judge outcomes in light of what they expected to happen. In this paper, we test this particular notion: whether expectations serve as a reference point.

A growing class of theories (e.g., David E. Bell 1985, Graham Loomes & Robert Sugden 1986, Faruk Gul 1991, Jonathan Shalev 2000, Botond Kőszegi & Matthew Rabin 2006, 2007, forthcoming) is built on the idea that expectations can act as a reference point. These models are able to align empirical evidence that is hard to reconcile with usual economic assumptions (e.g., Loomes & Sugden 1987, Paul Heidhues & Kőszegi 2008). Despite their theoretical and intuitive appeal, models of expectation-based reference-dependent preferences are inherently difficult to test, as expectations are hard to observe in the field. To sidestep this problem, we conduct a tightly controlled real-effort experiment. The two main advantages of our setup are that we know the rational expectations of participants about earnings and that we can exogenously influence these expectations. We are thus able to directly assess the relevance of theories of expectation-based reference-dependent preferences.

Investigating the importance of expectations helps with answering the key open question for reference-dependent preferences: what determines the reference point? Developing an empirically validated theory of where reference points come from is crucial for disciplining predictions. Otherwise, if the reference point is assumed case-by-case, models of reference-dependent preferences might explain behavior not because of their structural assumptions but because of this additional degree of freedom. Testing expectations as potential candidate for a reference point extends previous empirical research which has restricted attention mainly to the status quo or lagged status quo as reference point (e.g., Daniel Kahneman, Jack Knetsch, & Richard Thaler 1990, Terrance Odean 1998, David Genesove & Christopher Mayer

2001).

In our experiment, subjects work on a tedious and repetitive task. After each repetition, they decide whether to continue or to stop working. They get a piece rate, but receive their accumulated piece rate earnings only with 50 percent probability, whereas with 50 percent probability they receive a fixed, known payment instead. Which payment subjects receive is determined only *after* they have made their choice about when to stop working. The only treatment manipulation is a variation in the amount of the fixed payment.

By manipulating the amount of the fixed payment, we change the effort incentives of subjects with reference points in expectations. According to the models mentioned above, these individuals experience painful loss sensations if the realized state of the world compares unfavorably to a state that they expected could possibly happen instead. In our experiment, a subject can expect one of two states of the world to occur, with equal probability, when they stop working: receiving the fixed payment, f, or receiving piece rate earnings. Anticipating potential disappointment if they were to receive the less favorable of the two states, subjects can minimize potential loss sensations by stopping with piece rate earnings close to f. This way, no matter what happens, they know that unfavorable comparisons will not be too severe.[1] From the perspective of minimizing losses the best plan is actually to stop with earnings exactly equal to f; in this case a subject expects f in either state of the world, outcomes necessarily fulfill these expectations, and there is no chance of disappointment whatsoever. But whether it is optimal to completely eliminate potential loss, or instead just stop closer to f, depends on an individual's effort costs. Our treatment manipulation increases f and thus exogenously raises earnings expectations for the case that the fixed payment is realized. This means that an individual has to choose a higher effort level to reduce or eliminate a potential

---

[1]Note that while choosing to stop with piece rate earnings unequal f would allow for the possibility of a gain, if the more favorable outcome is realized, gains sensations are typically weaker than loss sensations (Kahneman & Amos Tversky 1979). Due to this asymmetry, known as "loss aversion", the subject would rather minimize deviations from expectations.

feeling of loss if they do not receive f. On the other hand they still do not want to work too far beyond f, to avoid being very disappointed in case they do receive f. Assuming a reference point in expectations thus predicts that increasing the size of the fixed payment will tend to increase overall effort, and that the propensity to stop is especially high when the piece rate earnings equal the fixed payment.

By contrast, a canonical model of effort provision with separable utility over money and effort costs does not predict a treatment difference. Optimal effort is determined by setting marginal cost equal to the marginal benefit defined by the piece rate, and the fixed payment is irrelevant for both marginal cost and marginal benefit. This is true independent of the shape of utility over money and the shape of the cost function, conditional on the assumption of separability. Models incorporating reference-dependent preferences but taking the status quo as the reference point also predict no treatment difference, because the status quo is the same across treatments.

Our data support the main predictions of reference-dependent preferences models with a reference point in expectations. When the amount of the fixed payment is large, subjects work significantly more than when the amount of the fixed payment is small. We also observe pronounced spikes in the distribution of effort choices, exactly at the low fixed payment amount in the low fixed payment treatment, and at the high fixed payment amount in the high fixed payment treatment. Moreover, there is no spike at the high fixed payment amount in the low fixed payment treatment, and vice versa. In additional control treatments, we show that our results are not driven by alternative, psychological mechanisms: subjects do not stop exactly at the fixed payment because this number is especially salient, and they do not work more when the fixed payment is higher because of reciprocal feelings towards the experimenter. Finally, we provide evidence that reference-dependent preferences are the key mechanism behind our results from another angle: We measure the degree of loss aversion of each subject by having them make a series of small-stakes lottery choices after the experiment. We find that subjects who are more loss averse according to this independent measure, stop closer to the fixed payment. This is a

unique prediction of the theory of a reference point in expectations, where stronger loss aversion leads to greater attraction to f.

One specific application of our findings is to the literature on labor supply and transitory wage changes. A series of studies have found evidence consistent with loss aversion around a daily reference income (e.g., Colin Camerer, Linda Babcock, George Loewenstein, & Thaler 1997, Yuan K. Chou 2002, Ernst Fehr & Lorenz Götte 2007, Henry S. Farber 2008, Vincent Crawford & Juanjuan Meng 2008, Kirk Bennett Doran 2009), with the exception of Farber (2005). In this literature the reference point has typically been treated as an unobserved, latent variable. Most closely related to our paper is the recent study by Crawford & Meng (2008) who use data on New York City taxi drivers' labor supply to test the theory of Kőszegi & Rabin (2006). They proxy the rational expectation about a driver's wage by the average wage earned per week day and find evidence for income and hours targeting around this expectation. Because there is no experimental variation, they address the problem of endogeneity using a structural approach. Our approach is complementary, in using a tightly controlled laboratory setting that allows us to exogenously vary rational expectations regarding earnings. Our studies find converging evidence on the importance of reference points in expectation for effort provision. We discuss the implications for the labor supply literature in more detail in Section 6.

Also related to our paper is the literature on violations of expected utility theory in lottery choices, in which some findings are supportive of a role for expectation-based reference points (see Loomes & Sugden (1987), Syngjoo Choi, Raymond Fisman, Douglas Gale & Shachar Kariv (2007) and Andreas Hack & Frauke Lammers (2008) for discussions). Different from our paper, this evidence has mainly come from inconsistencies observed in choices involving relatively complex combinations of different financial lotteries. Our experiment adds to this literature by measuring the impact of reference points as expectations in the domain of real effort choices, rather than lottery decisions. Moreover, it provides corroborating evidence on the importance of reference points as expectations, based on a simple and transparent test, where subjects can act in accordance with expected utility theory simply by

ignoring the fixed payment.

The paper is organized as follows. Details of the experimental design are explained in the following section. Section 3 discusses behavioral predictions. Results of the two main treatments are presented in Section 4. Section 5 reports on the control treatments and the lottery-based measure of loss aversion. Section 6 concludes.

2   Design

Our experiment was designed to create an environment that allows a precise measurement of behavior and in which we can exogenously influence a reference point in expectations. In the experiment, subjects worked on a tedious task. As the work task we chose counting the number of zeros in tables that consisted of 150 randomly ordered zeros and ones. This task does not require any prior knowledge, performance is easily measurable, and there is little learning possibility; at the same time, the task is boring and pointless and we can thus be confident that the task entailed a positive cost of effort for subjects. The task was also clearly artificial, and output was of no intrinsic value to the experimenter. This minimizes any tendency for subjects to use effort in the experiment as a way to reciprocate for payments offered by the experimenter.

The experiment involved two stages. Prior to the first stage, subjects read the instructions and answered control questions; they were also told that the experiment had a second stage but that details would be provided later.[2] During the first stage, subjects had four minutes to count as many tables as possible. They received a piece rate of 10 cents per correct answer for sure.[3] This part served to familiarize subjects with the task; due to this first stage, subjects had a good understanding of how

---

[2] An English translation of the instructions is provided in the appendix.

[3] In both stages, if an answer was not correct, subjects had two more tries for the same table. To prevent guessing, the piece rate was deducted from their account if they failed all three tries. This happened only 59 times in the experiment (compared to almost 12000 correctly counted tables).

difficult the task was and how much one could earn in a given time *before* they knew the amount of the fixed payment (which was revealed only after the first stage). Additionally, we will use performance in this stage as a productivity indicator.

After the first stage, subjects read the instructions for the second (and main) stage. The task was again to count zeros, but there were two differences compared to the first stage. First, they could now decide themselves how much and for how long they wanted to work. At most, they could work for 60 minutes. When they wanted to stop, they could push a button on the screen and the experiment was over: subjects answered a very short questionnaire (including a series of small-stakes lottery choices described in Section 5), got paid immediately, and could leave. How much subjects chose to work will be the main outcome variable in our analysis of the experiment. The second difference was that subjects did not get their accumulated piece rate earnings from the main stage for sure. Before they started counting in the main stage, they had to choose one of two closed envelopes. They knew that one of the envelopes contained a card saying "Acquired earnings" and that the other envelope contained a card saying "3 euros." But they did not know which card was in which envelope. The envelopes remained with the subjects while they were working and were only opened after the subject had stopped working. The subject's payment was then determined by the card in the chosen envelope. The piece rate per correct answer was doubled to 20 cents in the main stage in order to keep economic incentives comparable between the two stages.

We know the rational expectation of each subject regarding earnings, in the moment they were deciding whether to stop working: with 50 percent probability the subject would receive the accumulated piece rate earnings and with 50 percent he would receive the fixed payment. Because uncertainty about the payment was revealed only after the work was finished, we were able to exogenously vary a subject's rational expectation by changing the amount of the fixed payment.[4]

_____

[4]We don't know the *actual* expectations of subjects. The theories we are testing, however, all rely on the theoretical construct of rational expectations. Even so, it is unlikely that actual expectations are far from correct for many subjects in our setting: the lottery was very simple and salient and the potential payoffs (current accumulated piece rates and fixed payment) were always

7

There were two main treatments. The only difference between these treatments was the amount of the fixed payment. In the LO treatment, the fixed payment was 3 euros while it was 7 euros in the HI treatment. Treatments were assigned randomly to subjects; we also randomized treatments over morning and afternoon time slots and over days of the week.

A potential confound could have arisen if subjects worked in the same room and simultaneously started working, e.g., due to peer effects (Armin Falk & Andrea Ichino 2006) or due to a desire for conformity (Douglas Bernheim 1994). We employed a special procedure to prevent such effects: subjects arrived for the experiment one at a time, and individual starting times were at least 20 minutes apart. Upon arrival, subjects were guided to one of three essentially identical, neutral rooms.[5] They worked alone in their room with the door closed and never (with very few exceptions) saw another subject or the other two experimental rooms. Instructions and payments were also administered in their room. Because of this special procedure, subjects' stopping behavior could not have been influenced by other subjects' behavior in a systematic way.

We conducted three additional control treatments to check whether salience or reciprocity could have driven the results. Design and results of these treatments are described in Section 5.

Subjects were students from the University of Bonn studying various majors except Economics. We recruited subjects who had participated in no or only a few previous experiments. Experiments were computerized using the software z-Tree and ORSEE (Urs Fischbacher 2007, Ben Greiner 2004). 60 subjects participated in each treatment. No subject participated in more than one treatment. In addition to their earnings from the two stages of the experiment (on average 8.70 euros), subjects received a show-up fee of 5 euros. The experiment took about one hour on average, including time for instructions and both stages.

---

shown on the screen.

[5]Photos of the three rooms are shown in the appendix.

3 Predictions

We examine three categories of models: a canonical model with separable utility, models with status-quo reference dependence, and models with expectation-based reference dependence. Our setup can be described as follows: The subject's choice variable is the number of correctly solved tables e. With probability $\frac{1}{2}$ each, the subject receives either a fixed payment f or the accumulated piece rate earnings we, where $w > 0$ is the piece rate per table. $c(e)$ is the subject's cost of effort with $\partial c / \partial e > 0$ that the subject has to bear in both states of the world. Because we are interested in the effect of the size of f on effort provision, we set f to f_{LO} and f_{HI} for treatments LO and HI, respectively.[6]

First, consider a standard model of effort provision with a utility function separable in monetary payoff x and cost of effort: $U(x, e) = u(x) - c(e)$. In our setup, this utility function becomes $U(e, f, w) = \frac{1}{2}u(f) + \frac{1}{2}u(we) - c(e)$, yielding the following first-order condition:

$$\frac{\partial U}{\partial e} = \frac{w}{2}u'(we) - c'(e) \quad \Rightarrow \quad u'(we^*) = \frac{2}{w}c'(e^*)$$

The optimal effort level e^* is independent of the fixed payment f; if the subject receives the fixed payment f, he wishes to stop right away no matter how large f is. The prediction depends on the separability of money and cost of effort[7], but not

---

[6]As it happened only very rarely that a subject miscounted a table thrice and thus got the piece rate deducted from his earnings, we ignore this design detail in the predictions.

[7]It is common in labor economics to allow for non-separability across time periods in effort cost (e.g., fatigue) or in consumption utility (e.g., habit formation). With such forms of non-separability, the model still predicts no treatment difference. Adopting a less common assumption of contemporaneous non-separability of income and effort, e.g., with a function $U = g[u((we + f)/2) - c(e)]$, allows generating the prediction that effort increases with f, if the function $g[\cdot]$ is concave, i.e., if increasing expected wealth makes counting zeros less painful. The equally plausible assumption that having more money makes counting zeros more painful predicts that effort should decrease with f. Either way, such a utility function cannot explain the tendency to stop exactly at the fixed payment nor a correlation between individual loss aversion and stopping closer to f. Adding further assumptions to capture these other predictions becomes more and more ad hoc and

on the shape of the cost function nor on a subject's curvature in $u(\cdot)$, i.e., whether the subject is risk-neutral, risk-averse, or risk-loving.

Models incorporating reference-dependent preferences with the status quo as the reference point also predict no treatment difference, because the status quo when entering the experiment is the same across treatments and thus independent of f. The subject can affect how piece rate earnings compare to the status quo, but effort cannot influence the potential loss relative to status quo if the fixed payment is realized. Thus, varying the amount of the fixed payment has no impact on effort incentives, although the individual has reference-dependent preferences. A similar argument holds for reference points that may be affected by expectations about earnings that subjects had before learning about the exact incentive scheme for their particular treatment.

In contrast to these two models, theories assuming that agents have expectation-based reference-dependent preferences predict different behavior across treatments. Here, individuals dislike an outcome falling short of their expectations. We derive our hypotheses using the model of Kőszegi & Rabin (2007); models by Bell (1985), Loomes & Sugden (1986), and Gul (1991) generate similar predictions.[8] In Section 5, we derive further hypotheses for three control treatments and the lottery-based measure of individual loss aversion.

In Kőszegi & Rabin (2007), an individual derives "consumption utility" from the consumption bundle c and "gain-loss utility" from comparing c to a reference bundle r. Bundle r is the full distribution of rational expectations, i.e., every outcome that could have happened weighted with its ex-ante probability. As outcomes in our setup are not very large, we assume consumption utility to be linear and equal to c.

―――――――――――――

specific to our experimental setup, and makes it impossible to generalize the model to other settings. By contrast, the formulation of the reference-dependent preferences model below is inspired by a large body of empirical evidence and is easily generalized to different settings.

[8]The main difference between Kőszegi & Rabin (2007) and the other theories is how expectations are mapped into a reference point. Bell (1985), for example, assumes it to be the mean while Kőszegi & Rabin (2007) assume that an outcome is compared to the entire distribution of expectations; this distinction does not matter for our setup.

Overall utility is the sum of consumption and gain-loss utility, and is assumed to be separable across the K dimensions of c. We assume that subjects assess outcomes along 2 dimensions: money and effort costs. The gain-loss utility is defined by the function $\mu(c_k - r_k)$. For small arguments s, Kőszegi & Rabin assume that $\mu(s)$ is piece-wise linear: $\mu(s) = \eta s$ for $s \geq 0$ and $\mu(s) = \eta \lambda s$ for $s < 0$ with $\eta \geq 0$ and $\lambda > 1$; because λ is strictly greater than 1, losses loom larger than equal-sized gains.[9] We assume that the gain or loss sensation a subject finally experiences depends on their rational expectations about possible earnings amounts held the moment before the envelope is opened. The final piece rate earnings (and the fixed payment amount for the treatment) thus determine the reference point. This corresponds to the choice-acclimating equilibrium concept in Kőszegi & Rabin (2007), where the individual's choice shapes expectations, and thus the reference point that is held right before uncertainty is resolved. Because subjects pay the effort costs regardless of which envelope they draw, effort costs do not affect the comparison between the two states of the world and thus do not enter gain-loss utility.[10]

If the subject intends to stop at an accumulated earnings level below the fixed payment ($we < f$), the resulting expected utility will be given by

$$U = \frac{we + f}{2} - c(e) + \frac{1}{2}\eta \left[\frac{1}{2}(we - we) + \frac{1}{2}\lambda(we - f) \right] + \frac{1}{2}\eta \left[\frac{1}{2}(f - we) + \frac{1}{2}(f - f) \right]$$

The first two terms are expected consumption utility and cost of effort. The remaining terms are the expected gain-loss utility: the first bracketed term is the gain-loss utility when the outcome is we, multiplied by the probability of occurring ($\frac{1}{2}$) and by η, the strength of gain-loss utility. Inside this term, receiving we feels

---

[9]In its full generality, the model assumes that a stochastic outcome F is evaluated according to its expected utility, with the utility of each outcome being the average of how it feels relative to each possible realization of the reference point G: $U(F|G) = \int \int u(c|r) dG(r) dF(c)$. The reference point G is the probabilistic belief the individual held in the recent past about outcomes.

[10]Effort costs do still enter consumption utility. As for the canonical model, we make the simplifying assumption that subjects know their effort costs. One reason for the first stage of the experiment was to give subjects experience with their individual effort costs.

neutral; but receiving we while expecting the larger f feels like a loss. Since the subject expected to receive f with probability $\frac{1}{2}$, the terms are weighted accordingly. The second bracketed term shows gain-loss utility where the outcome is the fixed payment, applying the same logic.

If the accumulated earnings are higher than the fixed payment ($we \geq f$), the gain-loss utility is different. Receiving the accumulated earnings now feels like a gain compared to the lower fixed payment (third term), while receiving the fixed payment now means a loss (terms equal to zero are suppressed here):

$$U = \frac{we + f}{2} - c(e) + \frac{1}{2}\eta \left[\frac{1}{2}(we - f)\right] + \frac{1}{2}\eta \left[\frac{1}{2}\lambda(f - we)\right]$$

The first-order conditions are then:

$$we < f: \quad \frac{\partial U}{\partial e} = \frac{w}{2} - c'(e) + \frac{1}{4}\eta(\lambda - 1)w \quad \Rightarrow \quad c'(e^*) = \frac{w}{2} + \frac{w}{4}\eta(\lambda - 1)$$

$$we \geq f: \quad \frac{\partial U}{\partial e} = \frac{w}{2} - c'(e) - \frac{1}{4}\eta(\lambda - 1)w \quad \Rightarrow \quad c'(e^*) = \frac{w}{2} - \frac{w}{4}\eta(\lambda - 1)$$

When accumulated earnings are below f, the marginal returns to effort are higher than $\frac{w}{2}$, which is the return to effort in the canonical model without gain-loss utility (assuming linear $u(\cdot)$). Stopping entails a loss if the outcome turns out to be we rather than f; the pain of this loss more than offsets the potential pleasure of a gain if f is realized. When the accumulated earnings are above f, the incentive effect of loss aversion is reversed: because earnings beyond f can be lost in case the subject receives the fixed payment f, loss aversion now reduces the returns to effort relative to the canonical case. Gain-loss utility thus creates an additional incentive to exert effort when below the fixed payment amount, and reduces the incentive to work when above the fixed payment. Therefore, increasing the fixed payment should increase average effort, since it causes the marginal returns to remain high up to a higher effort level.

Hypothesis 1: *Average effort in the HI treatment is higher than in the LO treatment.*

Reference dependence moves optimal effort from above and below towards the fixed payment; the more loss averse a subject is the closer they should stop to the

fixed payment. For some subjects, this will even move optimal effort so far as to equalize expected piece rate earnings and f. The discrete drop in the return to effort at the fixed payment amount implies that there is a range of cost functions for which stopping exactly at the fixed payment is optimal. Thus, there will tend to be clustering of stopping decisions exactly at f. The stronger loss aversion is in the population, the larger the fraction stopping at f will be.

> **Hypothesis 2:** *The probability to stop at $we = f_{LO}$ is higher in the LO treatment than in the HI treatment; the probability to stop at $we = f_{HI}$ is higher in HI than in LO.*

We next turn to the empirical results from the experiment.

4   Results

Our first result supports Hypothesis 1. In the LO treatment with fixed payment $f = 3$ euros, subjects stop working after accumulating 7.37 euros on average. In the HI treatment with $f = 7$ euros, subjects stop on average at 9.22 euros.

> **Result 1:** *Subjects in the HI treatment work significantly more than subjects in the LO treatment.*

The treatment difference of 1.85 euros is almost half as large as the amount of the treatment manipulation ($7 - 3 = 4$ euros). The marginal effect compared to effort provision in LO is 25.1 percent. The treatment difference in effort provision is significant in an OLS regression where we compare effort in HI to effort in LO. We regress the accumulated earnings at which a subject stopped on a treatment dummy (see Table 1, column 1).[11] The treatment difference stays significant when we control for productivity, gender, outside temperature (experiments took place in the summer), and time of day. The only significant control variable is productivity

---

[11]The result is confirmed by non-parametric tests. A Mann-Whitney U-test yields a p-value of 0.015 (all p-values in this paper refer to two-sided tests). The same result obtains if we compare the distribution of stopping decisions: a two-sample Kolmogorov-Smirnov test rejects the equality of distributions between treatments ($p = 0.005$).

(Table 1, columns 2 and 3). As an indicator for productivity in the main stage, we use average time per correct answer in the first stage (measured in seconds multiplied by -1). A positive coefficient thus indicates that faster subjects complete more tables.[12]

It could be that the cost of effort is not only determined by the number of tables counted but also by the mere time subjects spend in the experiment. We therefore consider the time spent working as an alternative measure of effort provision. Treatments are also different for this dependent variable: subjects in LO work on average 31.7 minutes, while subjects in HI work on average 6.4 minutes longer, a marginal effect of 20.1 percent. This difference is significant in OLS regressions with and without the controls described above (see Table 1, columns 4 to 6).[13] Because subjects can only work between 0 and 60 minutes, we also present Tobit regressions that account for this censoring (Table 1, columns 7 to 9). This does not alter the results.[14]

As shown in Section 3, the model of Kőszegi & Rabin (2007) predicts that stopping decisions in the two treatments should differ in a very special way. Hypothesis 2 predicts a higher probability of stopping when the accumulated earnings equal the fixed payment. Neither the canonical model, nor the model with status quo reference

---

[12]The Spearman rank correlation coefficient between answering speeds in each stage is 0.520 ($p < 0.001$). This measure of productivity is not influenced by the treatment manipulation since subjects during the first stage did not know yet about the exact procedure of the main stage. Consequently, answering speed in the first stage is not significantly different between treatments (U-test, $p = 0.185$). Using average time per answer (i.e., including also wrong answers) or number of completed tables during the first stage instead of the measure used above does not change results.

[13]The treatment difference in working time is also statistically significant in non-parametric tests: U-test, $p = 0.034$; Kolmogorov-Smirnov test, $p = 0.085$.

[14]Censoring is not an issue if we take earnings as dependent variable; earnings are neither bounded above nor below (since subjects could make losses by miscounting tables thrice).

Table 1: Treatment Difference in Effort (HI compared to LO treatment)

	OLS: Accumulated earnings			OLS: Time spent working (in min.)			Tobit: Time spent working (in min.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1 if HI treatment	1.850**	1.942**	1.973**	6.430**	6.572**	6.784**	7.927**	8.091**	8.442**
	(0.917)	(0.885)	(0.900)	(3.163)	(3.153)	(3.231)	(3.841)	(3.814)	(3.833)
Productivity		0.059***	0.064***		0.091	0.096		0.098	0.103
		(0.019)	(0.020)		(0.067)	(0.070)		(0.080)	(0.083)
1 if Female			-0.039			1.619			1.577
			(0.950)			(3.412)			(4.035)
Controls for temperature	No	No	Yes	No	No	Yes	No	No	Yes
Controls for time of day	No	No	Yes	No	No	Yes	No	No	Yes
Constant	7.370***	10.607***	10.200***	31.715***	36.713***	34.362***	33.004***	38.389***	35.306***
	(0.648)	(1.206)	(1.445)	(2.237)	(4.297)	(5.190)	(2.697)	(5.143)	(6.116)
N.Obs.	120	120	120	120	120	120	120	120	120
Adjusted or Pseudo R^2	0.03	0.09	0.08	0.03	0.03	0.00	0.00	0.01	0.01

Notes: The dependent variable is the level of accumulated earnings (in euro) at which a subject stopped working for columns 1–3, and time spent working (in minutes) until a subject stopped for columns 4–9. Columns 1–6 report results from OLS regressions, columns 7–9 show results of Tobit regressions (the lower and upper limits are 0 and 60 minutes). Data from LO and HI treatments are included in the analysis. The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Adjusted R^2 is shown for OLS; pseudo R^2 for Tobit. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

dependence, make this prediction. Our data are consistent with Hypothesis 2.

> **Result 2:** *The probability to stop when accumulated earnings are equal to the amount of the fixed payment is higher compared to the same earnings level in the other treatment. The modal choice in both treatments is to stop exactly when accumulated earnings equal the fixed payment.*

Figure 1 shows a histogram of accumulated earnings (LO in the top panel, HI in the bottom panel). First of all, stopping decisions are dispersed over a wide range. Some subjects stop directly, others work for up to 25 euros. This is what one would expect given that productivity and cost of effort differ across subjects. But there are systematic differences between treatments in terms of clustering of stopping decisions exactly at the fixed payment: in the LO treatment, many subjects stop at 3 euros (15.0 percent of subjects); in the HI treatment, almost nobody stops at 3 euros (1.7 percent). By contrast, in HI many subjects stop at 7 euros (16.7 percent); in LO very few subjects stop here (3.3 percent). The modal choice in both treatments is to stop exactly when accumulated earnings equal the fixed payment. These treatment differences are statistically significant. Results of a multinomial logit regression with the three outcomes "stop at 3 euros", "stop at 7 euros", and "stop elsewhere" are presented in Table 2. Column 1 shows the regression without controls, in columns 2 and 3 the controls used in Table 1 are added. Being in the HI treatment leads to significantly less stopping at 3 euros and more stopping at 7 euros compared to being in the LO treatment.[15] The same results obtain if we compare the number of subjects stopping in a range around 3 and 7 euros. For example between 2 and 4 euros, 30.0 and 5.0 percent of subjects stop in LO and HI, respectively (U-test, $p < 0.001$); between 6 and 8 euros, these figures are 13.3 and 36.7 percent, respectively (U-test, $p = 0.003$). Multinomial logit estimates for this result are presented in Table 6 in the appendix.

---

[15]These differences are also significant in non-parametric tests: the percentages of subjects stopping at 3 euros is significantly higher in LO (U-test, $p = 0.009$); the percentage stopping at 7 euros is higher in HI (U-test, $p = 0.015$).
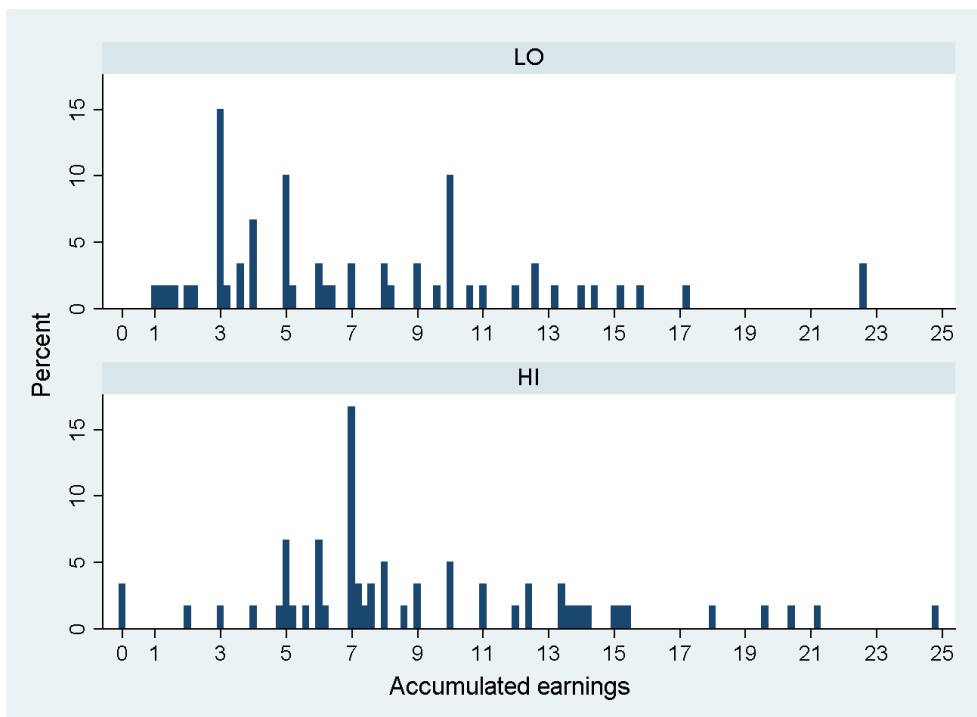
Figure 1: Histogram of Accumulated Earnings (in Euro) at Which a Subject Stopped.

Table 2: Tendency to Stop at the Fixed Payment (HI compared to LO treatment)

	(1a) Stop at 3	(1b) Stop at 7	(2a) Stop at 3	(2b) Stop at 7	(3a) Stop at 3	(3b) Stop at 7
1 if HI treatment	-2.197** (1.073)	1.609** (0.801)	-2.191** (1.074)	1.620** (0.802)	-2.318** (1.115)	1.781** (0.829)
Productivity			0.003 (0.014)	0.005 (0.016)	-0.003 (0.019)	0.004 (0.020)
1 if Female					-1.094 (0.789)	0.106 (0.661)
Controls for temperature	No		No		Yes	
Controls for time of day	No		No		Yes	
Constant	-1.695*** (0.363)	-3.199*** (0.721)	-1.523* (0.848)	-2.946*** (1.121)	-1.437 (1.215)	-3.032** (1.326)
N.Obs.	120		120		120	
Pseudo R^2	0.09		0.09		0.17	

Notes: The table reports results of multinomial logit regressions. The dependent variable indicates three outcomes: "stop at 3 euros", "stop at 7 euros", and "stop elsewhere" which is the reference category. Data from LO and HI treatments are included in the analysis. The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

5   Robustness Checks

In the previous section we presented evidence supporting models of expectation-based reference-dependent preferences: subjects work more when expectations are high, and many subjects stop when piece rate earnings equal the fixed payment, thus avoiding any potential loss relative to expectations.

In this section, we check whether other, psychological motivations could also have played a role in generating the observed treatment differences. We first present results of three control treatments showing that neither salience nor reciprocity significantly influences effort provision in our setting. We then provide further, direct evidence that loss aversion is a key mechanism driving our findings: We use an independent measure of an individual's degree of loss aversion, and show that more loss averse subjects stop closer to the fixed payment.

5.1   Salience

It is conceivable that stopping decisions did not cluster at the fixed payment because of reference dependence but because the fixed payment was salient. If subjects resorted to irrelevant, environmental cues to decide when to stop, they might have stopped at 3 euros or 7 euros because these amounts were mentioned frequently in the instructions and also on the computer screens and could have served as a focal point.[16] Similarly, it could be that the salience of the fixed payment influenced subjects to set arbitrary goals of earning these amounts (Edwin A. Locke & Gary P. Latham 1990, 2002) and thus impacted behavior through this channel rather than reference-dependent preferences.[17]

---

[16]Focal points or arbitrary anchors have been shown to influence behavior by, e.g., Tversky & Kahneman (1974), Karen Jacowitz & Kahneman (1995), Gretchen Chapman & Eric Johnson (1999), and David K. Whynes, Zoë Philips, & Emma Frew (2005).

[17]In a recent paper, Alexander K. Koch & Julia Nafziger (2008) demonstrate that this intuition could also go in the other direction. They build upon the model of Kőszegi & Rabin (2006) to show that goal setting can be explained by assuming that individuals have expectation-based reference-dependent preferences (see also Chip Heath, Richard Larrick, & George Wu 1999).

To check whether the salience of, e.g., "3 euros" influences behavior in our experiment, we conducted two symmetric treatments, both variants of the LO treatment. The NOSAL treatment (NOSAL for *no salience*) suppresses the salience of "3 euros" as much as possible but keeps the reference-dependence motive to stop at 3. The SAL treatment (SAL for *salience*) keeps the salience of the fixed payment exactly the same as in the LO treatment but removes the loss aversion motive to stop there. The general procedure of the two treatments was identical to the LO treatment: subjects came one-by-one, they counted zeros in tables after choosing one of two envelopes, and the card in their chosen envelope determined their payoff when they decided to stop working. We again collected 60 observations per treatment. The only difference was the two cards. Recall that in LO, the two cards read "Acquired earnings" and "3 euros."

In the NOSAL treatment, the cards read "5 euros plus acquired earnings" and "8 euros." Thus, unlike in LO, the number "3" was never mentioned in the description of the potential payoffs or on the computer screens and could therefore not have acted as a convenient anchor. Still, the way to avoid potential losses was to stop at 3 euros. The additional 5 euros were taken out of the show-up fee, so total expected earnings were identical to the LO treatment; the two treatments differed merely in the framing of payoffs. If salience is important in our setting, behavior in NOSAL should differ from behavior in LO: there should be no special tendency to stop at 3 in NOSAL. Subjects should rather stop more often at 5 or 8 euros than subjects in LO. As the salient numbers in NOSAL point to larger amounts compared to LO, average effort should be higher than in LO. Expectation-based reference-dependent preferences, however, predict that behavior in NOSAL and LO should not be different and that also in NOSAL subjects should be especially likely to stop exactly at 3.[18]

In the SAL treatment, the two cards read "Acquired earnings" and "Acquired

---

[18]The canonical model of effort provision and a model of loss aversion around the status quo similarly predict no difference between NOSAL and LO. But at the same time these models do not predict that there should be clustering of stopping decisions at 3 in either treatment.

earnings plus 3 euros." This means that subjects in SAL actually received the accumulated piece rate for sure and played an additional lottery (0, 3 euros; 0.5). To keep incentives for a rational, risk neutral subject the same as in LO, the piece rate in SAL was halved to 10 cents (since subjects received the piece rate only with 50 percent probability in LO but got it for sure in SAL). Salience of "3 euros" remained exactly as in the LO treatment: every occurrence of "3 euros" in the original instructions or screens was replaced by the phrase "acquired earnings plus 3 euros" where applicable. "3 euros" was thus mentioned equally often and at the same places as in the LO treatment.

If the high probability of stopping at 3 is driven by salience, there should be a tendency to stop right at 3 in SAL, as was the case in LO. But if subjects have expectation-based reference-dependent preferences, the treatments should differ, with no clustering of stopping decisions at 3 in SAL. Expected utility in SAL, according to the model by Kőszegi & Rabin (2007), is given by:

$$U = \frac{we + f}{2} - c(e) + \frac{1}{2}\eta\left[\frac{1}{2}f\right] + \frac{1}{2}\eta\left[\frac{1}{2}\lambda(-f)\right]$$

If the subject receives f in addition to the piece rate earnings, this feels like a gain relative to the alternative of only getting the piece rate earnings (third term). If the subject only receives the piece rate earnings, this feels like a loss relative to also getting the additional amount f (fourth term). The loss of f is weighted more heavily than the gain of f, leading to a lump-sum reduction in utility, but this does not influence optimal effort. In contrast to the main treatments, subjects in SAL cannot influence the size of a potential loss by choosing a particular effort level. Therefore, unlike in the LO treatment, expectation-based loss aversion does not predict a tendency for stopping decisions to cluster at $f = 3$, because this does not help avoid losses. The probability to stop at $we = f$ should thus be lower in SAL than in LO.[19]

---

[19] The standard model of effort provision with a separable, linear utility function predicts no difference between LO and SAL. In such a model, SAL implies exactly the same incentives as LO: $U = (we + f)/2 - c(e)$, with $f = 3$. Reference dependence around the status quo also predicts no

We designed SAL primarily to test the different predictions of reference-dependent preferences and salience regarding when subjects will, and will not, cluster stopping decisions exactly at 3 euros. One can also compare average effort levels in SAL and LO, but in this case the models make more similar predictions. The reference-dependent preferences model predicts that effort in SAL will be higher than in LO. As shown in Figure 1, most subjects in LO stopped above the fixed payment. For these subjects, loss aversion held back their effort. In SAL, the loss aversion motive to hold back effort is removed, as there is no risk of feeling a loss when getting the piece rates, so subjects should work harder, potentially the full time. If salience—and not reference dependence—drives behavior, then there should be clustering of stopping decisions at 3 in SAL, which requires subjects to count more tables than in LO and thus raises average effort (the nominal piece rate was reduced to 10c as subjects got the piece rate for sure).[20]

Data from the two control treatments do not support the salience explanation for clustering of stopping decisions at the fixed payment amount:

> **Result 3:** *The probability to stop exactly at 3 euros is not significantly different between NOSAL and LO, despite the difference in salience; and subjects stop less often at 3 euros in SAL compared to LO, although salience is held constant.*

In the NOSAL treatment, 13.3 percent of subjects stop at 3 euros compared to 15.0 percent in the LO treatment. In SAL, only 3.3 percent of subjects stop at 3 euros. Results of multinomial logit regressions comparing stopping behavior in NOSAL and SAL to LO are shown in Table 3.[21] The dependent variable is whether subjects stopped at 3 euros, at 7 euros, or somewhere else. Compared to

---

treatment difference.

[20]The canonical model with linear utility predicts the same effort level in SAL and LO. But if utility is concave, a sure piece rate of 10c will be valued higher than a 20c piece rate with 50 percent probability. In that case, effort in SAL should again be higher.

[21]In addition, the sample contains observations from our reciprocity treatment (R treatment) which we discuss in more detail below. The results also hold in separate regressions where we restrict the sample to only LO and NOSAL or to LO and SAL.

the LO treatment, stopping behavior is not different in NOSAL. Subjects in SAL, however, stop significantly less often at 3 euros. These results continue to hold when we include the control variables used in Tables 1 and 2 (columns 2 and 3) or if we consider stopping in a range around 3 euros, see Table 7 in the appendix.[22] In addition, subjects in NOSAL do not stop more often at 5 or 8 euros than subjects in LO, showing that just mentioning a number saliently does not create a tendency to stop there in our setting.[23] Moreover, subjects in NOSAL stop more often at 3 euros than subjects in HI (U-test, $p = 0.016$) while this is not true for SAL (U-test, $p = 0.560$). These results suggest that subjects in the main treatments do not stop at f because of salience. In contrast, the results are consistent with the model assuming expectations-based reference dependence: People stop at 3 when this helps minimize losses, even if 3 is minimally salient, and do not exhibit a special tendency to stop at 3 if stopping there cannot help avoid losses, although 3 is salient.

Data on average effort levels are also in line with predictions of expectation-based reference-dependent preferences:

> **Result 4:** *There is no difference between NOSAL and LO in average effort provision. Subjects in SAL work significantly more than subjects in LO.*

Table 4 shows OLS and Tobit estimates of average effort regressed on treatment dummies without and with the controls described above. Accumulated earnings when stopping are not significantly different between NOSAL and LO (columns 1–3) although salient numbers (5 and 8) are higher than the salient number 3 in

---

[22]The SAL-LO difference of stopping at 3 euros is also significant in non-parametric tests (U-test, $p = 0.027$) while the NOSAL-LO difference is not significant (U-test, $p = 0.794$). A potential concern could arise because subjects in SAL had to complete 30 tables to reach the fixed payment of 3 euros while subjects in LO needed only 15 tables. If effort costs simply made it impossible to reach 30 tables in SAL, this would mechanically prevent subjects from stopping exactly at 3 euros. However, 65 percent of subjects in SAL completed at least 30 tables but only the above mentioned 3.3 percent stopped exactly at 30.

[23]1.7 percent of NOSAL-subjects and 3.3 percent of LO-subjects stop at 8 euros (U-test: $p = 0.560$). In both treatments, 10 percent of subjects stop at 5 euros.

Table 3: Tendency to Stop at the Fixed Payment (Control Treatments Compared to LO)

	(1a) Stop at 3	(1b) Stop at 7	(2a) Stop at 3	(2b) Stop at 7	(3a) Stop at 3	(3b) Stop at 7
1 if SAL treatment	-1.638** (0.806)	-0.134 (1.019)	-1.619** (0.807)	-0.117 (1.020)	-1.989** (0.902)	-0.259 (1.147)
1 if NOSAL treatment	-0.076 (0.527)	0.958 (0.861)	-0.081 (0.528)	0.954 (0.862)	-0.052 (0.591)	1.341 (0.931)
1 if R treatment	-0.666 (0.592)	-0.078 (1.019)	-0.584 (0.599)	-0.009 (1.027)	-0.605 (0.688)	0.368 (1.088)
Productivity			0.008 (0.010)	0.007 (0.013)	0.009 (0.010)	0.007 (0.014)
1 if Female					-0.411 (0.455)	-0.533 (0.657)
Controls for temperature	No		No		Yes	
Controls for time of day	No		No		Yes	
Constant	-1.695*** (0.363)	-3.199*** (0.721)	-1.258** (0.620)	-2.838*** (0.996)	-1.004 (0.756)	-3.076*** (1.193)
N.Obs.	240		240		240	
Pseudo R^2	0.04		0.04		0.07	

Notes: The table reports results of multinomial logit regressions. The dependent variable indicates three outcomes: "stop at 3 euros", "stop at 7 euros", and "stop elsewhere" which is the reference category. Data from LO, SAL, NOSAL, and R treatments are included in the analysis. The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

LO. As the nominal piece rate in SAL differs, it is misleading to compare earnings between SAL and LO. We thus take the number of correctly completed tables as measure of effort in columns 4–6. Subjects in SAL work significantly more than in LO; this is consistent with loss aversion motives having held back effort in LO. This also holds in Tobit regressions that account for the fact that the number of correctly solved tables cannot be negative (columns 7–9).[24] Taking time worked as alternative measure of effort yields a similar result: interestingly, subjects in NOSAL work slightly shorter (U-test, $p = 0.095$) while subjects in SAL work significantly longer than in LO (U-test, $p < 0.001$). 30.0 percent of subjects in SAL work even the full 60 minutes compared to 13.3 percent in LO and 10.0 percent in NOSAL. Finally, the treatment difference in effort between HI and NOSAL is highly significant (e.g., accumulated earnings: U-test, $p < 0.001$).

5.2 Reciprocity

Another potential motive, distinct from salience, could be reciprocity (Rabin 1993, Falk & Fischbacher 2006). Numerous studies have shown that many individuals reward kind acts even if it is not in their immediate interest (for an overview see Fehr & Simon Gächter 2000). It might thus be that subjects in HI worked more than in LO because they perceived the higher average pay in HI as a kind act by the experimenter and reciprocated by counting more tables. Note, however, that reciprocity (unlike reference-dependent preferences) cannot explain why many subjects stop exactly at the fixed payment; reciprocity could potentially explain only the result on average effort.

While reciprocity has been shown to be important in many contexts, it is likely

---

[24]The SAL-LO difference is also significant in non-parametric tests (tables completed: U-test, $p = 0.005$; Kolmogorov-Smirnov test, $p = 0.018$) while the NOSAL-LO difference is not significant (accumulated earnings: U-test, $p = 0.310$; Kolmogorov-Smirnov test, $p = 0.432$; tables completed: U-test, $p = 0.305$; Kolmogorov-Smirnov test, $p = 0.432$). The results also hold in separate OLS regressions where we restrict the sample to only LO and NOSAL or to LO and SAL.

Table 4: Treatment Difference in Effort (Control Treatments Compared to LO)

	OLS: Accumulated earnings			OLS: Tables completed			Tobit: Tables completed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1 if SAL treatment				12.417***	12.878***	15.429***	12.417***	12.897***	15.403***
				(4.468)	(4.424)	(4.778)	(4.522)	(4.467)	(4.772)
1 if NOSAL treatment	-1.183	-1.195	-1.119	-6.017	-6.118	-5.955	-6.816	-6.923	-6.541
	(0.857)	(0.856)	(0.942)	(4.468)	(4.420)	(4.841)	(4.536)	(4.477)	(4.841)
1 if R treatment	-1.173	-1.004	-1.269	-5.933	-4.432	-5.809	-6.330	-4.780	-5.995
	(0.857)	(0.867)	(0.983)	(4.468)	(4.461)	(5.000)	(4.529)	(4.511)	(5.005)
Productivity		0.014	0.014		0.123**	0.130***		0.128**	0.136***
		(0.011)	(0.011)		(0.050)	(0.050)		(0.050)	(0.050)
1 if Female			-0.311			-2.264			-2.137
			(0.709)			(3.145)			(3.148)
Controls for temperature	No	No	Yes	No	No	Yes	No	No	Yes
Controls for time of day	No	No	Yes	No	No	Yes	No	No	Yes
Constant	7.370***	8.135***	9.455***	37.050***	43.825***	51.045***	37.050***	44.100***	51.633***
	(0.606)	(0.869)	(1.090)	(3.159)	(4.155)	(4.989)	(3.198)	(4.199)	(4.992)
N.Obs.	180	180	180	240	240	240	240	240	240
Adjusted or Pseudo R^2	0.00	0.01	0.01	0.08	0.10	0.11	0.01	0.01	0.02

Notes: The dependent variable is the level of accumulated earnings (in euro) at which a subject stopped working for columns 1–3, and the number of tables completed correctly for columns 4–9. Columns 1–6 report results from OLS regressions, columns 7–9 show results of Tobit regressions (the lower limit is 0 tables). Data from LO, SAL, NOSAL, and R treatments are included in the analysis (SAL only for columns 4–9). The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Adjusted R^2 is shown for OLS; pseudo R^2 for Tobit. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

to be irrelevant in our setting for several reasons.[25] Nevertheless, we conducted an additional control treatment to test whether the level of total pay triggers a reciprocal reaction. The R treatment (R for *reciprocity*) is identical to the LO treatment except for an additional lump sum payment of 4 euros which is announced at the beginning of the main stage and is clearly linked to working on the task (the instructions read "For the work in this part of the experiment, you receive 4 euros up front."). This raises total expected earnings relative to the LO treatment and even surpasses expected earnings in the HI treatment.[26] If the higher effort in HI compared to LO was simply due to a reciprocal reaction to a lump-sum payment of 7 rather than 3, then average effort should also be higher in R than in LO. If, however, reference-dependent preferences drove the treatment difference, behavior in R should not be different from LO.

Result 5: *Effort provision in R and LO is not significantly different.*

Table 4 shows OLS regressions where we regress average effort measured by earnings or completed tables on treatment dummies (LO treatment is the omitted category) and various controls described above. We find that average effort is not significantly different in R compared to LO. If anything, the point estimates suggest a reduction in effort.[27] Table 3 additionally shows that clustering of stopping decisions

---

[25]First, previous studies show a role for reciprocity primarily between subjects, rather than between subjects and the experimenter where social distance is probably larger (George A. Akerlof 1997, Gary Charness, Ernan Haruvy, & Doron Sonsino 2007); second, direct tests of whether subjects feel reciprocal towards the experimenter conclude that this is in fact not the case (Björn Frank 1998); third, counting zeros clearly has no intrinsic value to the experimenter and thus working hard does not benefit the experimenter; fourth, there is suggestive evidence that piece rate compensation tends to cue income maximization, rather than reciprocity (Stephen Burks, Jeffrey Carpenter, & Götte 2009); fifth, we strongly emphasized in the instructions that subjects were free to work as little or as much as they wanted.

[26]We chose an additional payment of 4 euros to make sure that, even if subjects disregard the 50 percent probability of getting the fixed payment in HI and just focus on the "nominal" earnings, perceived earnings in R are still as high as in HI.

[27]The result holds also in non-parametric tests (accumulated earnings: U-test, $p = 0.178$;

in R is also not significantly different from LO. Moreover, the treatment difference in effort between HI and R is highly significant (e.g., accumulated earnings: U-test, $p < 0.001$).[28] We therefore conclude that reciprocity did not drive the difference between the LO and the HI treatment.

5.3 Evidence on Individual Loss Aversion and Stopping Decisions

After each of our experiments, subjects answered a short questionnaire. Among other questions, we asked subjects to state reasons for their stopping decision. Answers were given in free form without any suggestion of possible reasons. Of those subjects stopping exactly when accumulated earnings equal the fixed payment, the great majority named reasons such as a desire to avoid disappointment if they drew the less favorable envelope, or that they wanted to "make sure" to get the amount of the fixed payment by working at least that much. Because they indicated a preference to avoid unfavorable comparisons to what might have happened, we interpret these answers as providing evidence that reference dependence and disappointment aversion were important for generating clustering of stopping at the fixed payment.

But it would be more desirable to have an incentivized measure of subjects' individual loss aversion to directly investigate the link between strength of loss aversion on the one hand and stopping behavior on the other. During the questionnaire, we therefore had subjects make six choices, each time between a fixed payment of 0 and a small-stakes lottery. The lottery involved a 50/50 chance of winning 6 euros

Kolmogorov-Smirnov, $p = 0.587$; tables completed: U-test, $p = 0.180$; Kolmogorov-Smirnov, $p = 0.587$) or if we take time worked as dependent variable (U-test, $p = 0.250$; Kolmogorov-Smirnov, $p = 0.432$).

[28] If we pool the three treatments that are equivalent according to models of reference-dependent preferences (LO, NOSAL, and R) and regress accumulated earnings or time spent working on a treatment dummy for HI (as in Table 1), the p-values of the treatment coefficient all drop below 0.01, the highest p-value being 0.003. Mann-Whitney U-tests of these pooled treatment differences yield p-values below 0.001.

or receiving Y. Across lotteries, Y was varied from -2 to -7 euros in steps of 1 euro. Subjects knew that one of the six choices would be randomly selected and, if they had chosen the lottery, this lottery would be played out for money. Note that the small stakes mean that rejections of lotteries with positive expected value cannot be explained by standard risk aversion (Rabin 2000). Rather, the number of lotteries that a subject rejects gives an indicator for the individual's degree of loss aversion. A very similar measure has been used in previous studies, and has been shown to predict loss-averse behavior in terms of labor supply, as well as strength of the endowment effect (Fehr & Götte 2007, Gächter, Andreas Herrmann, & Johnson 2007). We thus take the number of rejected lotteries as a proxy for the individual's degree of loss aversion. For our setting it does not matter whether this proxy concerns the strength of the gain-loss utility, η, or the loss aversion parameter, λ.[29]

If subjects have no reference-dependent preferences, we should not expect subjects to reject small-stakes lotteries with positive expected value. But even if they did, there should be no correlation between the number of rejected lotteries and the stopping decision in the experiment (as long as an individual's effort cost is uncorrelated with their risk aversion) because the stopping decision is only determined by an individual's effort cost function. Status-quo-based loss aversion also predicts no correlation with stopping behavior, nor do salience or reciprocity explanations. Additionally, subjects learned about the lotteries only after they stopped working; thus the anticipation of the lotteries could not have influenced stopping behavior directly.

If subjects have expectations-based reference-dependent preferences, however, we should expect a positive correlation between the number of rejected lotteries and the stopping decision: the more loss averse a subject is, the closer to f they should

---

[29]The number of rejected lotteries increases in the degree of loss aversion regardless of the definition of the reference point. This is clearly true when the status quo of zero is the reference point. It is also true for an expectation-based reference point: Using Kőszegi & Rabin's (2007) choice acclimating personal equilibrium, as we did to derive the predictions in Section 3, there is a cutoff value for Y below which rejecting the lottery is the preferred personal equilibrium. The cutoff increases in η and λ, i.e., a more loss averse subject rejects more lotteries.

stop. To see this, remember that the first-order condition for earning levels below f is $c'(e^*) = \frac{w}{2} + \frac{w}{4}\eta(\lambda - 1)$. If η or λ increase, i.e., if a subject gets more loss averse, the optimal effort increases and thus moves closer to f. For earning levels above f, the condition is $c'(e^*) = \frac{w}{2} - \frac{w}{4}\eta(\lambda - 1)$; an increase in η or λ decreases optimal effort, moving it closer to f.[30] Our data support this hypothesis:

Result 6: *Subjects who are more loss averse stop significantly closer to f.*

Table 5 shows results of OLS regressions in which we regress the absolute distance of accumulated earnings to the fixed payment on a subject's degree of loss aversion and various controls. Data from all four treatments where reference-dependent preferences could impact stopping behavior (LO, HI, NOSAL, R) are included.[31] We find in all specifications that the proxy for loss aversion is significantly negative, i.e., more loss-averse subjects stop closer to the fixed payment.[32] This provides additional, direct support for reference-dependent preferences and loss aversion as a key

---

[30]The model by Kőszegi & Rabin (2007) also makes a more subtle prediction, which could generate this correlation for reasons related to endogenous background risk. If subjects stop far from f they could try to "hedge" the resulting risk by accepting more lotteries. Since the lotteries are independent from the draw of the envelope, this can (in some cases) reduce a potential loss. This prediction relies on special assumptions about choice bracketing, however, that are not very plausible. Namely we would need to assume that subjects do not consider risk from outside the experiment but that they do consider risk from the main stage of the experiment when deciding on the lotteries. Numerous studies (e.g., Tversky & Kahneman 1981, Rabin & Georg Weizsäcker forthcoming) show that subjects usually bracket more narrowly, and do not even bracket together two different lotteries that are shown on the same decision sheet. Either way, only Kőszegi & Rabin's (2007) model predicts this correlation and other explanations (e.g., salience or non-reference dependent risk aversion) do not.

[31]2 out of 240 subjects chose lotteries inconsistently: they switched more than once between the safe and the risky option which makes it difficult to interpret their choices. These two subjects are excluded from the analysis. Results are unchanged if we include them.

[32]In SAL we expect there to be no relationship between loss aversion and stopping at 3, given that stopping at 3 does not help avoid losses. Indeed, if we estimate the same regressions as in Table 5 for SAL, without controls the point estimate is close to zero (-0.041) and far from significant. Adding controls, the effect gets closer to zero, and is even positive in the final specification.

mechanism in the main treatment difference.

6 Conclusion

In a simple real-effort laboratory experiment, we tested theories of reference-dependent preferences that assume the reference point to be a function of individual expectations. Our data is in line with predictions of these models: subjects work more when expectations are high, and many subjects stop when piece rate earnings equal the fixed payment. Three additional treatments ruled out alternative explanations based on salience and reciprocity. We also provided direct evidence from lottery choices that reference-dependent preferences drive our results. Our results thus contribute to understanding what determines the reference point. They support models which assume the reference point to be formed by expectations, like Bell (1985), Loomes & Sugden (1986), Gul (1991), or Kőszegi & Rabin (2006).

Our results are also relevant for the literature on reference points and labor supply. Studies in this literature use field data on worker effort choices to test whether the response of effort to changes in incentives is consistent with the standard intertemporal substitution of labor and leisure or rather with loss aversion around a daily reference income. In this literature the reference point has typically been treated as an unobserved, latent variable. Camerer et al. (1997) demonstrated that the daily labor supply of NYC cab drivers is in line with loss aversion around a daily income target. Camerer et al. (1997) and Farber (2005) both note, however, that daily earnings vary too much to be explained by a *fixed* daily income target. Partly in response to this evidence, Kőszegi & Rabin (2006) developed a theory of expectation-based reference-dependent preferences that allows the income target to differ in a predictable way across days. Our experiment adds to this literature by making the rational expectations about earnings known to the researcher and by providing exogenous variation while keeping other potential reference points constant. As noted by Kőszegi & Rabin (2006) and subsequently shown by Crawford & Meng (2008) using the data set of Farber (2005), if reference points are based on

Table 5: Impact of Loss Aversion on Distance to Fixed Payment

| | $|we - f|$ | | | |
	(1)	(2)	(3)	(4)
Loss aversion	-0.489**	-0.500**	-0.518**	-0.472**
	(0.220)	(0.222)	(0.222)	(0.236)
Productivity			0.013	0.014
			(0.009)	(0.010)
1 if Female				-0.188
				(0.578)
Controls for treatments	No	Yes	Yes	Yes
Controls for temperature	No	No	No	Yes
Controls for time of day	No	No	No	Yes
Constant	6.040***	6.726***	7.522***	7.368***
	(0.934)	(1.050)	(1.191)	(1.273)
N.Obs.	238	238	238	238
Adjusted R^2	0.02	0.01	0.02	0.01

Notes: The table reports estimates from OLS regressions. The dependent variable $|we - f|$ is the absolute distance of accumulated earnings to the fixed payment (in euro). Data from LO, HI, NOSAL, and R treatments are included in the analysis. Two subjects who chose inconsistently in the lottery measure are excluded. The proxy for loss aversion is the number of lotteries that a subject rejected (see text for details). The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

expectations, *anticipated* changes in incentives should not distort behavior relative to standard theory, given that expectations adjust to reflect the anticipated change. For example, if an individual expects the hourly wage to be low on a given day, earning a small amount does not feel like a loss. But if the hourly wage is *unexpectedly* low, this does feel like a loss relative to expectation, and can induce workers to work even harder to try to reach their expectation, contrary to the standard prediction on intertemporal substitution which implies that workers should decrease effort when the wage is temporarily low. This distinction helps reconcile some of the seemingly conflicting findings in the field evidence. Our results are complementary, providing controlled evidence that expectations can in fact act as a reference point, and can affect effort provision.

An interesting direction for future research is to distinguish between different expectation-based models of reference-dependent preferences. Our treatments are not designed to test which way of specifying the reference point in expectations is the empirically most plausible: assuming that the reference point is the mean of the expected outcomes (like in Bell 1985, Loomes & Sugden 1986, or Gul 1991) or assuming that the reference point is the whole distribution of expected outcomes (like in, e.g., Kőszegi & Rabin 2006, 2007). Both of these assumptions predict a higher probability to stop when accumulated earnings equal the fixed payment. Our experimental design provides a useful platform for pursuing this question in the future, however, and could be extended to distinguish between these models: if subjects' final payoffs are determined by a lottery over *two* distinct fixed payments and accumulated earnings, rather than just one fixed payment and accumulated earnings as in the current study, then predictions are different across models. Models like the one of Loomes & Sugden (1986) predict a higher probability to stop when accumulated earnings equal the mean of the two fixed payments but not when they equal one of the two fixed payments. Models like Kőszegi & Rabin (2006) predict a higher probability to stop at the two fixed payments but not at the mean.

7 References

Akerlof, G. A. (1997), 'Social distance and social decisions', *Econometrica* **65**(5), 1005–1027.

Bell, D. E. (1985), 'Disappointment in decision making under uncertainty', *Operations Research* **33**(1), 1–27.

Bernheim, D. (1994), 'A theory of conformity', *Journal of Political Economy* **102**(5), 842–877.

Burks, S., Carpenter, J. & Götte, L. (2009), 'Performance pay and the erosion of worker cooperation: Field experimental evidence', *Journal of Economic Behavior & Organization* **70**(3), 459–469.

Camerer, C., Babcock, L., Loewenstein, G. & Thaler, R. (1997), 'Labor supply of New York City cabdrivers: One day at a time', *Quarterly Journal of Economics* **112**(2), 407–41.

Chapman, G. & Johnson, E. (1999), 'Anchoring, activation, and the construction of values', *Organizational Behavior and Human Decision Processes* **79**, 115–153.

Charness, G., Haruvy, E. & Sonsino, D. (2007), 'Social distance and reciprocity: An internet experiment', *Journal of Economic Behavior & Organization* **63**(1), 88–103.

Choi, S., Fisman, R., Gale, D. & Kariv, S. (2007), 'Consistency and heterogeneity of individual behavior under uncertainty', *American Economic Review* **97**(5), 1921–1938.

Chou, Y. K. (2002), 'Testing alternative models of labor supply: Evidence from cab drivers in Singapore', *The Singapore Economic Review* **47**(1), 17–47.

Crawford, V. & Meng, J. (2008), 'New york city cabdrivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income', *UCSD Discussion Paper* .

Doran, K. B. (2009), 'The existence and position of daily income reference points: Implications for daily labor supply', *University of Notre Dame Discussion Paper* .

Falk, A. & Fischbacher, U. (2006), 'A theory of reciprocity', *Games and Economic Behavior* **54**(2), 293–315.

Falk, A. & Ichino, A. (2006), 'Clean evidence on peer effects', *Journal of Labor Economics* **24**(1), 39–57.

Farber, H. S. (2005), 'Is tomorrow another day? The labor supply of New York City cabdrivers', *Journal of Political Economy* **113**(1), 46–82.

Farber, H. S. (2008), 'Reference-dependent preferences and labor supply: The case of New York City taxi drivers', *American Economic Review* **98**(3), 1069–1082.

Fehr, E. & Gächter, S. (2000), 'Fairness and Retaliation: The Economics of Reciprocity', *Journal of Economic Perspectives* **14**, 159–181.

Fehr, E. & Götte, L. (2007), 'Do workers work more when wages are high? Evidence from a randomized field experiment', *American Economic Review* **97**(1), 298–317.

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

Frank, B. (1998), 'Good news for experimenters: subjects do not care about your welfare', *Economics Letters* **61**(2), 171–174.

Gächter, S., Herrmann, A. & Johnson, E. J. (2007), 'Individual-level loss aversion in riskless and risky choices', *University of Nottingham Discussion Paper* .

Genesove, D. & Mayer, C. (2001), 'Loss aversion and seller behavior: Evidence from the housing market', *Quarterly Journal of Economics* **116**(4), 1233–1260.

Greiner, B. (2004), 'An online recruitment system for economic experiments', *Forschung und wissenschaftliches Rechnen* **63**, 79–93.

Gul, F. (1991), 'A theory of disappointment aversion', *Econometrica* **95**(3), 667–686.

Hack, A. & Lammers, F. (2008), 'The role of expectations in the formation of reference points over time', *unpublished* .

Heath, C., Larrick, R. & Wu, G. (1999), 'Goals as reference points', *Cognitive Psychology* **38**, 79–109.

Heidhues, P. & Kőszegi, B. (2008), 'Competition and price variation when consumers are loss averse', *American Economic Review* **98**(4), 1245–1268.

Jacowitz, K. & Kahneman, D. (1995), 'Measures of anchoring in estimation tasks', *Personality and Social Psychology Bulletin* **21**(11), 1161–1170.

Kahneman, D., Knetsch, J. & Thaler, R. (1990), 'Experimental tests of the endowment effect and the Coase theorem', *Journal of Political Economy* **98**, 1325–1348.

Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**, 263–291.

Kőszegi, B. & Rabin, M. (2006), 'A model of reference-dependent preferences', *Quarterly Journal of Economics* **121**(4), 1133–1165.

Kőszegi, B. & Rabin, M. (2007), 'Reference-dependent risk attitudes', *American Economic Review* **97**(4), 1047–1073.

Kőszegi, B. & Rabin, M. (forthcoming), 'Reference-dependent consumption plans', *American Economic Review* .

Koch, A. K. & Nafziger, J. (2008), 'Commitment to self-rewards', *IZA Discussion Paper* .

Locke, E. A. & Latham, G. P. (1990), *A theory of goal setting & task performance*, Prentice-Hall.

Locke, E. A. & Latham, G. P. (2002), 'Building a practically useful theory of goal setting and task motivation.', *American Psychologist* **57**(9), 705–717.

Loomes, G. & Sugden, R. (1986), 'Disappointment and dynamic consistency in choice under uncertainty', *The Review of Economic Studies* **53**(2), 271–282.

Loomes, G. & Sugden, R. (1987), 'Testing for regret and disappointment in choice under uncertainty', *Economic Journal* **97**, 118–129.

Odean, T. (1998), 'Are investors reluctant to realize their losses?', *Journal of Finance* **53**(5), 1775–1798.

Rabin, M. (1993), 'Incorporating fairness into game theory and economics', *American Economic Review* **83**(5), 1281–1302.

Rabin, M. (2000), 'Risk aversion and expected-utility theory: A calibration theorem', *Econometrica* **68**(5), 1281–1292.

Rabin, M. & Weizsäcker, G. (forthcoming), 'Narrow bracketing and dominated choices', *American Economic Review* .

Shalev, J. (2000), 'Loss aversion equilibrium', *International Journal of Game Theory* **29**(2), 269–287.

Tversky, A. & Kahneman, D. (1974), 'Judgment under uncertainty: Heuristics and biases', *Science* **185**(4157), 1124–1131.

Tversky, A. & Kahneman, D. (1981), 'The framing of decisions and the psychology of choice', *Science* **211**, 453–458.

Whynes, D. K., Philips, Z. & Frew, E. (2005), 'Think of a number ... any number?', *Health Economics* **14**, 1191–1195.

A   Instructions

Below are the instructions of the LO treatment translated into English. The numbers in brackets mark places where changes in the instructions for the other treatments are made. These changes are described in detail after the instructions.

The experiment consists of two parts. Please start by reading the explanations for the first part carefully. You will receive the instructions for the second part of the experiment after the first part is finished.

[1] For your arrival on time, you receive **5 euros** that will be paid to you at the end of the experiment. If you have any questions during the experiment please ask the experimenter. If you use the computer in an improper way you will be excluded from the experiment and from any payment.

Instructions for the first part of the experiment

What do you have to do?

In this part of the experiment your task is to count zeros in a series of tables. The figure shows the work screen you will use later:
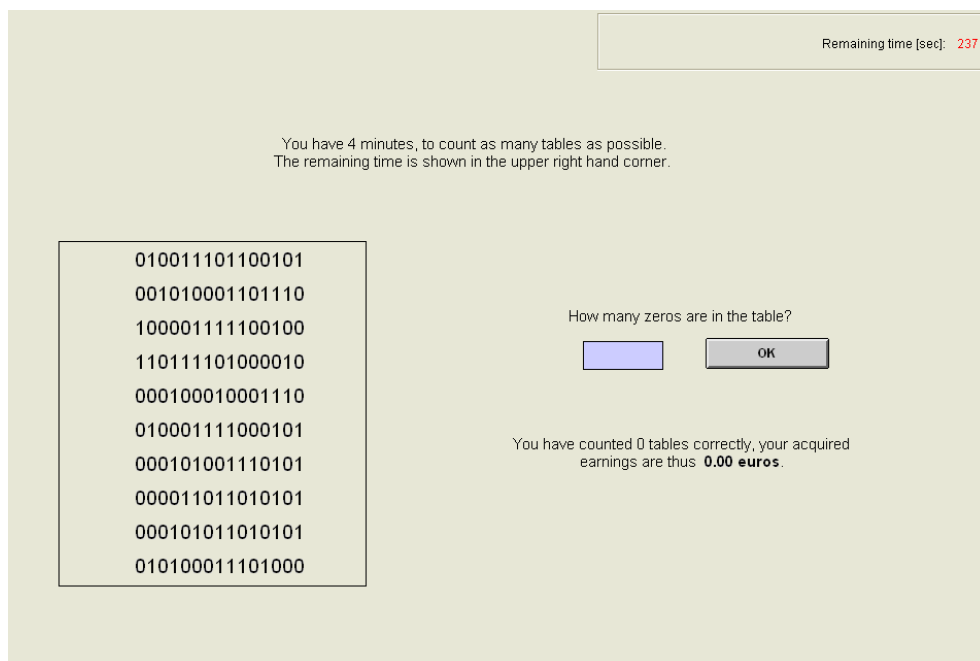
Enter the number of zeros into the box on the right side of the screen. After you have entered the number, click the OK-button. If you enter the correct result, a new table will be generated. If your input was wrong, you have two additional tries

to enter the correct number into the table. You therefore have a total of three tries to solve each table.

If you entered the correct number of zeroes you earn money: **You receive 10 cents for each table you solved correctly.**

If you enter three times a wrong number for a table, 10 cents will be subtracted from your earnings and a new table will then be generated. The earnings of this part of the experiment will be paid to you at the end of the experiment.

Example: You solve three tables correctly; you miscount one table once. You miscount a fourth table three times. Your earnings are therefore:

> 3 x 10c for the correctly counted tables
>
> - 1 x 10c for the fourth table, which you miscounted three times.

thus a total of 20c.

You have 4 minutes until the first part of the experiment is over. The remaining time is displayed in the upper right hand corner of the screen.

Counting tips: Of course you can count the zeros any way you want. Speaking from experience, however, it is helpful to always count two zeros at once and multiply the resulting number by two at the end. In addition you miscount less frequently if you track the number you are currently counting with the mouse cursor.

Example question

Please answer the following question:

Assume you have solved 5 tables correctly, and miscounted two tables three times.

> What are your acquired earnings?   _____ euros

After you have answered the example question correctly, the experimenter will start the first part of the experiment.

Instructions for the second part of the experiment

What do you have to do?

[2] The task in this part of the experiment is once again to count zeros in a series of tables. The figure shows the work screen you will use later:
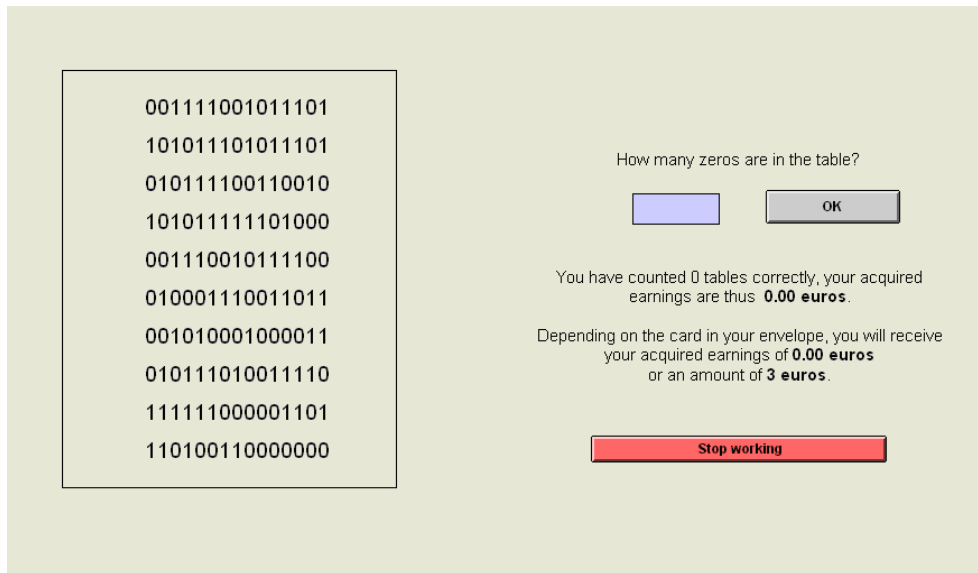
[3] New rules are now in effect, which did not apply in the first part:

- [4] **For each correctly solved table you will be credited 20 cents.** After three wrong inputs 20 cents will be subtracted from your earnings.

- [5] It is possible to lose the acquired earnings from this part of the experiment: there are two envelopes in front of you. One envelope contains a card with the text "acquired earnings", the other contains a card "3 euros". **You do not know which card is in which envelope.** Please choose one of the envelopes now and sign on the envelope.

- While you are working, the envelopes will remain in your room. After you have finished your task, we will open the envelopes. **You are not allowed to open the envelopes** before you have finished your task and one of the experimenters is with you.

- **If you have drawn the envelope with the card "acquired earnings", you will get your acquired earnings and not the 3 euros.**

- [6] **If you have drawn the envelope with the card "3 euros", you will get 3 euros and not the acquired earnings.** The amount of 3 euros does not change, no matter how many tables you solved.

- After your work is done we will also open the envelope which you did not choose, such that you can check that the envelopes contained different cards.

Important: In this part of the experiment you can count zeros as long as you want. This means you can decide yourself when you want to stop working. You can work, however, at most 60 minutes.

If you want to stop counting, please click on the **red button** "stop working" and contact us by briefly stepping into the corridor. You will be paid in your room.

Example: You stop after ten minutes and have solved 24 tables correctly with no miscounts. Your acquired earnings are therefore $24 \times 20c = 4.80$ euros. The envelope chosen by you contains the card "acquired earnings". You therefore get 4.80 euros.

Example: You stop after 10 minutes and have solved 24 tables correctly with no miscounts. Your acquired earnings are therefore 4.80 euros. The envelope chosen by you contains the card "3 euros". You therefore get 3 euros instead of the 4.80 euros.

[7] **Example:** You stop after 30 minutes and have solved 4 tables correctly and miscounted three times at a 5th table. Your acquired earnings are therefore $4 \times 20c - 1 \times 20c = 60c$. The envelope chosen by you contains the card "3 euros". You therefore get the amount of 3 euros instead of your acquired earnings of 60c.

Example questions

Please answer the following questions:

Assume you have solved 28 tables correctly within 20 minutes.

- What are your acquired earnings?   _____euros

- How much money do you get if the envelope chosen by you contains the card "acquired earnings"?   _____euros

- How much money do you get if the envelope chosen by you contains the card "3 euros"?   _____euros

Assume you have solved 7 tables correctly within 15 minutes.

- What are your acquired earnings?   _____euros

- How much money do you get if the envelope chosen by you contains the card "acquired earnings"?   _____euros

- How much money do you get if the envelope chosen by you contains the card "3 euros"?   _____euros

After you have answered the example questions correctly, the experimenter will start the second part of the experiment.

The instructions for the other treatments differ as follows:

- *HI treatment: Every occurrence of "3 euros" is replaced by "7 euros" (also in the screenshot).*

- *NOSAL treatment: In paragraph [1], the first sentence regarding the show-up fee is dropped. Every occurrence of "3 euros" is replaced by "8 euros" and every occurrence of "acquired earnings" is replaced by "5 euros plus the acquired earnings" (also in the screenshot). In paragraph [6], the last sentence is changed to "The amounts of 5 and 8 euros do not change, no matter how many tables you count." The example calculations are adapted according to the payoff rules of NOSAL.*

- *SAL treatment: Every occurrence of "3 euros" is replaced by "acquired earnings plus 3 euros" where applicable (also in the screenshot). Before paragraph [3], a new paragraph is inserted reading: "For each correctly solved table you will again be credited 10 cents. After three wrong inputs 10 cents will be subtracted from your earnings." Paragraph [4] is correspondingly dropped. In paragraph [5], the first sentenced up to the colon is dropped. The example calculations are adapted according to the payoff rules of SAL.*

- *R treatment: At the beginning of paragraph [2], a sentence is added "For the work in this part of the experiment, you receive 4 euros up front." After paragraph [7], a paragraph is added: "You will receive the 4 euros up front (and all earnings from the fist stage of the experiment) in any case in addition."*

B Photos of Experimental Rooms
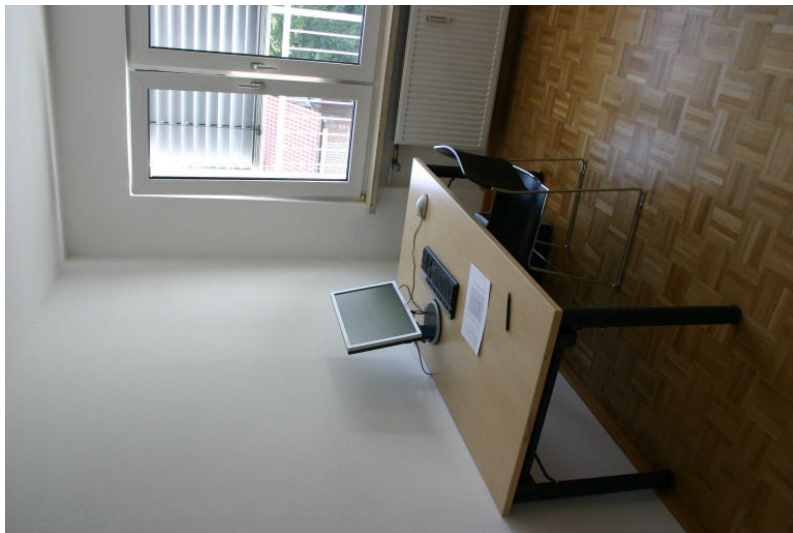
C Additional Regression Tables

Table 6: Tendency to Stop Around the Fixed Payment (HI compared to LO treatment)

	(1a) Stop in 2–4	(1b) Stop in 6–8	(2a) Stop in 2–4	(2b) Stop in 6–8	(3a) Stop in 2–4	(3b) Stop in 6–8
1 if HI treatment	-1.792*** (0.669)	1.056** (0.477)	-1.807*** (0.671)	1.062** (0.478)	-1.983*** (0.707)	1.109** (0.497)
Productivity			-0.007 (0.009)	0.002 (0.011)	-0.010 (0.011)	0.002 (0.011)
1 if Female					-0.168 (0.574)	0.567 (0.491)
Controls for temperature	No		No		Yes	
Controls for time of day	No		No		Yes	
Constant	-0.636** (0.291)	-1.447*** (0.393)	-1.033* (0.606)	-1.354* (0.698)	-1.170 (0.792)	-1.484* (0.803)
N.Obs.	120		120		120	
Pseudo R^2	0.08		0.09		0.14	

Notes: The table reports results of multinomial logit regressions. The dependent variable indicates three outcomes: "stop between 2 and 4 euros", "stop between 6 and 8 euros", and "stop elsewhere" which is the reference category. Data from LO and HI treatments are included in the analysis. The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

Table 7: Tendency to Stop Around the Fixed Payment (Control Treatments Compared to LO)

	(1a) Stop in 2–4	(1b) Stop in 6–8	(2a) Stop in 2–4	(2b) Stop in 6–8	(3a) Stop in 2–4	(3b) Stop in 6–8
1 if NOSAL treatment	-0.463	0.348	-0.465	0.341	-0.728	0.448
	(0.443)	(0.515)	(0.443)	(0.519)	(0.502)	(0.593)
1 if R treatment	0.198	0.210	0.252	0.375	-0.024	0.544
	(0.409)	(0.546)	(0.416)	(0.557)	(0.498)	(0.622)
Productivity			0.004	0.017	0.005	0.017
			(0.006)	(0.011)	(0.006)	(0.011)
1 if Female					0.119	0.452
					(0.365)	(0.434)
Controls for temperature	No		No		Yes	
Controls for time of day	No		No		Yes	
Constant	-0.636**	-1.447***	-0.404	-0.577	-0.405	-1.022
	(0.291)	(0.393)	(0.428)	(0.680)	(0.552)	(0.792)
N.Obs.	180		180		180	
Pseudo R^2	0.01		0.02		0.05	

Notes: The table reports results of multinomial logit regressions. The dependent variable indicates three outcomes: "stop between 2 and 4 euros", "stop between 6 and 8 euros", and "stop elsewhere" which is the reference category. Data from LO, NOSAL, and R treatments are included in the analysis. As the nominal piece rate in SAL is halved, stopping between 2 and 4 euros includes a wider range of stopping decisions. We thus exclude this treatment from this analysis. The proxy for productivity is the time subjects needed per table during the first stage (in seconds multiplied by -1). Standard errors are in parentheses. Significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.