DISCUSSION PAPER SERIES

IZA DP No. 3939

Reference Points and Effort Provision

Johannes Abeler
Armin Falk
Lorenz Götte
David Huffman

January 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Reference Points and Effort Provision

Johannes Abeler
University of Nottingham and IZA

Armin Falk
University of Bonn and IZA

Lorenz Götte
University of Geneva and IZA

David Huffman
Swarthmore College and IZA

Discussion Paper No. 3939
January 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3939
January 2009

ABSTRACT

Reference Points and Effort Provision*

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, risk aversion, disappointment, experiment

Corresponding author:

Johannes Abeler
University of Nottingham
School of Economics
University Park
Nottingham, NG7 2RD
United Kingdom
E-mail: johannes.abeler@nottingham.ac.uk

* Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged. We thank Steffen Altmann, Roland Benabou, Vincent Crawford, Lucie Dörner, Uri Gneezy, Paul Heidhues, Oliver Kirchkamp, Botond Kőszegi, Dorothea Kübler, Juanjuan Meng, Susanne Ohlendorf, Andrew Oswald, Matthew Rabin, Sam Schulhofer-Wohl, Klaas Schulze, Barry Schwartz, Daniel Seidmann, Chris Starmer, Kostas Tatsiramos, Georg Weizsäcker, and Matthias Wibral for helpful discussions. Valuable comments were also received from seminar and conference participants in Amsterdam, Bonn, Cologne, Gummersbach, Jena, London, Lyon, Milan, Nottingham, Trento, and Zurich. Anke Becker, Benedikt Vogt, and Ulf Zoelitz provided excellent research assistance.

1 Introduction

Imagine two identical workers. One expected a salary increase of 10 percent but received an increase of only 5 percent. The other received the same 5 percent wage increase but had not expected an increase. The change in income was the same for both workers, but the first worker probably 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., Bell 1985, Loomes & Sugden 1986, Gul 1991, Shalev 2000, Kőszegi & Rabin 2006, 2007, forthcoming) is built on the idea that expectations can act as a reference point and these models are able to align empirical evidence that is hard to reconcile with usual economic assumptions (e.g., Loomes & Sugden 1987, Heidhues & Kőszegi 2008). But 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 and that we can exogenously vary 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., Kahneman et al. 1990, Odean 1998, Genesove & Mayer 2001).

In our experiment, subjects work on a tedious and repetitive task. After each repetition, they can 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 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 changing the fixed payment, we vary earnings expectations of our subjects. Therefore, expectation-based reference-dependent preferences models predict effort provision to differ between treatments. Intuitively, if a subject's accumulated piece rate earnings are lower than the fixed payment, receiving the piece rates falls short of his expectations, as he partly expected to get the (higher) fixed payment. Loss aversion[1] relative to expectations thus provides an additional incentive for an individual to work hard, because this closes the gap between possible outcomes and expectations. Once the accumulated piece rate earnings are higher than the fixed payment, the incentive effect of loss aversion is reversed: now, the subject might end up with the (lower) fixed payment and this possibility *reduces* the incentive to exert effort. Reference dependence therefore moves the optimal effort towards the effort level that equalizes accumulated piece rates and the fixed payment. Models that assume a reference point in expectations thus predict 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 of effort. This is true independent of the shape of utility over money and the shape of the cost function, conditional on the assumption of separability.

[1]Loss aversion means that people dislike an outcome falling short of the reference point—a "loss"—more than they like an outcome exceeding the reference point—a "gain" (Kahneman & Tversky 1979).

Models incorporating reference-dependent preferences but taking the status quo as the reference point also predict no treatment difference because the status quo when entering the experiment is the same for both treatments.

We find a significant treatment effect, such that individuals in the high fixed payment treatment work more. The size of the increase in effort provision is large relative to the treatment manipulation: average earnings increase by about 2 euros in response to a 4 euro difference in the fixed payment amount across treatments. 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 an additional control treatment, we show that our results are not driven by a focal-point effect: subjects do not stop at the fixed payment anymore when we keep the salience of the fixed payment constant but move the reference point away from it. Taken together, these findings support the main predictions of reference-dependent preferences models with a reference point in expectations.

One specific application of our findings is to the lively dispute regarding labor supply and transitory wage changes. There is an ongoing controversy whether the response of labor supply 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 (e.g., Camerer et al. 1997, Farber 2005, Fehr & Götte 2007, Farber 2008, Doran 2007, Crawford & Meng 2008). 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 observe expectations directly and vary them exogenously. 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 5.

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 Bell (1985), Loomes & Sugden (1986, 1987), and Hack & 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 are presented in Section 4. Section 5 concludes.

2 Design

Our experiment is designed to create an environment that allows measuring behavior and expectations and in which we can exogenously vary rational 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 eliminates any tendency for subjects to use effort in the experiment as a way to reciprocate for any incentives offered by the experimenter.

The experiment involved two stages. Prior to the first stage, subjects read the instructions and answered control questions.[2] They knew that a second stage would follow but had no more details about the second stage. 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 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 got paid immediately, answered a very short questionnaire, 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 rates of 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

[2]For an English translation of the instructions, see Appendix A.

[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 45 times in the experiment (vs. almost 8000 correctly counted tables).

keep economic incentives comparable between the two stages.

We know the rational expectation of each subject when 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.

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 (Falk & Ichino 2006) or due to a desire for conformity (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.[4] 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 also conducted an additional control treatment to check whether the salience of "3 euros" and "7 euros" could have driven behavior in the two main treatments. It could be that subjects pay attention to environmental cues to decide when to stop.[5] "3 euros" and "7 euros" were mentioned several times in the instructions

[4]Photos of the three rooms are shown in Appendix B.

[5]Focal points or arbitrary anchors have been shown to influence behavior by, e.g., Tversky & Kahneman (1974), Mehta et al. (1994), Jacowitz & Kahneman (1995), Chapman & Johnson

and on the computer screen and could thus have served as a focal point. The procedure of the control treatment (called *SAL* for salience) 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. The only difference was the two cards. In the LO-treatment, the two cards read "Acquired earnings" and "3 euros." In the SAL-treatment, however, the cards read "Acquired earnings" and "Acquired 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 the LO-treatment, the piece rate in the SAL-treatment was reduced to 10 cents (since subjects received the piece rate only with 50 percent probability in the LO-treatment but got it for sure in the SAL-treatment). 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.

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 (Fischbacher 2007). 180 subjects participated in the experiment, 60 in each treatment. No subject participated in more than one treatment. In addition to their earnings from the two stages of the experiment, subjects received a show-up fee of 5 euros. On average, subjects earned 13.70 euros (about 18.40 USD at the time of the experiment). The experiment took about one hour on average, including time for instructions and both stages.

(1999), and Whynes et al. (2005).

3 Predictions

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

First, consider a standard model of effort provision with a utility function separable in monetary payoff x and the 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 because in the state of the world where the subject receives the fixed payment f, he wants to stop right away, independent of the amount of the fixed payment. And in the state of the world where he gets the accumulated earnings, the fixed payment does not matter either. Although linear utility would be the most reasonable assumption for the stake sizes in the experiment, this result also holds if the subject is risk-averse, or risk-loving. The shape of the cost function also has no influence. This prediction, however, depends on the separability of money and cost of effort.[7]

[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] Lifting the assumption of separability, it is conceivable that some or all of the predictions listed below can be derived without assuming reference dependence explicitly. It is not straightforward,

Models incorporating reference-dependent preferences but taking the status quo as the reference point also predict no treatment difference. The reason is that the status quo when entering the experiment is the same for both treatments and thus independent of f. A similar argument holds for any other reference point that is constant across treatments. This includes lagged expectations, i.e., expectations about earnings that subjects might have had, before showing up for the experiment and learning about the actual incentive scheme for their particular treatment.

In contrast to predictions made by the two previous models, theories in which agents have expectation-based reference-dependent preferences predict different behavior across treatments. In these models, individuals compare outcomes to their expectation and dislike an outcome falling short of expectations. We derive four testable hypotheses using the model of Kőszegi & Rabin (2006, 2007), but the models by Bell (1985), Loomes & Sugden (1986), and Gul (1991) generate similar predictions.[8]

In the model of Kőszegi & Rabin (2006), an individual derives "consumption utility" from the level of the consumption bundle c. As the outcomes in our setup are not very large, we assume the consumption utility to be linear, $m(c) = c$. Additionally, the individual derives "gain-loss utility" from comparing c to a reference bundle r. Overall utility is the sum of consumption and gain-loss utility. Both consumption utility and gain-loss utility are assumed to be separable across K dimensions. The gain-loss utility is defined by $n(c_k|r_k) = \mu(c_k - r_k)$, i.e., a function of the difference between the intrinsic consumption utilities of the realized outcome and the reference outcome. 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 $\lambda > 1$. The fact that λ is strictly greater than 1 captures loss aversion: losses loom larger than

however, to construct such a model which is at the same time not completely ad hoc. We therefore don't pursue this possibility further.

[8]The main difference between Kőszegi & Rabin (2006) 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 (2006) assume that an outcome is compared to the entire distribution of expectations; but this distinction does not matter for our setup.

equal-sized gains. Since we conduct a real effort experiment, it is natural to assume that subjects assess outcomes along $K = 2$ dimensions: money and effort/leisure. The reference outcome r is the *full distribution of rational expectations*, i.e., every outcome that could have happened weighted with its ex-ante probability. Gain-loss utility in a given state of the world is thus determined by comparing the outcome that happened to each of the outcomes that could have happened.[9]

If current accumulated earnings are below the fixed payment ($we < f$) and the subject decides to stop, the resulting expected utility is 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: expected earnings, and the cost of effort the subject has to bear in either state of the world. The remaining, bracketed terms are the expected gain-loss utility. The first term in brackets is the gain-loss utility in the case that the outcome is we. The whole bracket is multiplied by $\frac{1}{2}$, the probability that this outcome actually occurs and by η, the strength of gain-loss utility. Receiving we is compared to we, generating zero gain-loss sensations; but receiving we while expecting the larger f feels like a loss. That term is thus multiplied by $\lambda > 1$. Since the subject expected to receive f with probability $\frac{1}{2}$, the term is weighted accordingly. The second bracketed term shows gain-loss utility in the case that the outcome is the fixed payment. There is zero gain-loss utility from comparing f to itself, but receiving f feels like a gain compared to the lower we. The cost of effort never shows up in the gain-loss utility as the subject decides after each table whether to continue or to stop working. Thus the expected cost of effort

[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 beliefs the individual held in the recent past about outcomes. We calculate what Kőszegi & Rabin (2007) call the "choice-acclimating personal equilibrium", since subjects had time to adjust their expectations to the particular incentives faced in the experiment when they made their decisions to stop working.

is always the same as the actual cost. There is only uncertainty about the monetary payoff.

If the current accumulated earnings are higher than the fixed payment ($we \geq f$), the gain-loss utility changes. 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 in this equation):

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

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 $u(\cdot)$). This is because stopping would entail 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, so there is an unambiguous increase in the return to effort. When the accumulated earnings are above f, the incentive effect of loss aversion is reversed: loss aversion now reduces the returns to effort relative to the canonical case, as earnings beyond f can be lost in case the subject receives the fixed payment f. 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.*

The discrete drop in the return to effort at the fixed payment amount implies that there is a whole range of cost functions for which stopping exactly at the fixed payment is optimal. Thus, the model predicts clustering of stopping decisions exactly at the fixed payment f. The model does not predict that *every* subject stops exactly at the fixed payment. The probability of subjects stopping will be higher when accumulated earnings equal the fixed payment, but the percentage of subjects stopping at this amount depends on individuals' cost of effort and the strength of the potential reference dependence of each individual.

> **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 the HI-treatment than in the LO-treatment.*

It is conceivable that subjects do not stop at the fixed payment because of reference dependence but because the fixed payment is salient. If subjects resort to uninformative, environmental cues to decide when to stop, they might stop at 3 euros or 7 euros because these amounts are mentioned frequently in the instructions and also on the computer screens and could serve as a focal point. To exclude this alternative explanation, we turn to the SAL-treatment, which keeps the salience of the fixed payment exactly the same as in the LO-treatment but moves the reference point away from it. If behavior in the SAL-treatment is different from the LO-treatment, we can conclude that the treatment difference between LO- and HI-treatment is due to reference dependence and not due to salience.

In the SAL-treatment, subjects received either $(\frac{1}{2}w)e$, or $(\frac{1}{2}w)e + f_{LO}$ with equal probability (the piece rate was half that in the LO-treatment). In a canonical model of effort provision with a separable, linear utility function, the SAL-treatment implies exactly the same incentives as the LO-treatment: $U = (we + f)/2 - c(e)$. Such a model therefore predicts behavior to be the same in SAL- and LO-treatment. Reference dependence around the status quo also predicts no treatment difference.

If subjects have expectation-based reference-dependent preferences, as in the model by Kőszegi & Rabin (2006), the expected utility in SAL is given by

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

The first two terms are expected earnings and expected cost of effort. The third and fourth terms show expected gain-loss utility. If the subject receives f_{LO} 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_{LO} (fourth term).

Total utility now is lower than if the individual did not have reference-dependent preferences: the loss of f_{LO} is weighted more heavily than the gain of f_{LO}, leading to a lump-sum reduction in utility. But the first order condition, $c'(e) = \frac{w}{2}$, is the same as for the canonical model with linear utility. 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, the model does not predict a tendency for stopping decisions to cluster at f_{LO}; the return to effort does not depend on being below or above f_{LO}.

> **Hypothesis 3:** *The probability to stop at $we = f_{LO}$ is higher in the LO-treatment than in the SAL-treatment.*

The prediction for how average effort in SAL and LO compare depends on whether the fixed payment in the LO-treatment holds back effort for most subjects or induces more effort for most subjects. Individuals who would stop working at an earnings level below f_{LO} in the LO-treatment would work even less hard in the SAL-treatment, because without the motive of avoiding a loss the return to effort is lower. Individuals who would stop with earnings greater than f_{LO} in the LO-treatment would work harder in the SAL-treatment, as loss aversion held back their effort in the LO-treatment; in the SAL-treatment, they can work harder without the risk of feeling a loss when getting the piece rates.[10] If we assume that the average

[10]Subjects who stopped exactly at f_{LO} in the LO-treatment could work more or less in the SAL-treatment depending on whether loss aversion reduced or increased their effort in LO.

absolute impact of loss aversion on individual effort is the same above and below f we can make the following prediction:

> **Hypothesis 4:** *If most subjects in the LO-treatment stop below f_{LO}, effort provision is lower in the SAL-treatment compared to the LO-treatment. If most subjects in the LO-treatment work for more than f_{LO}, average effort provision should be higher in the SAL-treatment.*

In summary, the model with reference dependence in expectations generates four testable predictions regarding treatment differences. A canonical model of effort provision with a separable utility function predicts effort to be the same across the three treatments; status-quo reference dependence also predicts no treatment difference.

4 Results

In this section we present results of the experiment. The main variable of interest is the accumulated earnings at which a subject decided to stop. We start by comparing the two main treatments, HI and LO, testing Hypotheses 1 and 2; we then turn to the analysis of the third treatment, SAL, focusing on Hypotheses 3 and 4.

The first result supports Hypothesis 1:

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

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. The treatment difference of 1.85 euros is quite large, almost half the amount of the treatment manipulation ($7 - 3 = 4$ euros). The marginal effect compared to effort provision in the LO-treatment is 25.1 percent. The treatment difference in effort provision is significant in an OLS regression where 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 (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. Figure 1 shows that the answering speed is very stable between the two stages, consistent with performance in the first stage reflecting a stable productivity characteristic.[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 the LO-treatment work on average 31.7 minutes, while subjects in the HI-treatment 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

[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 ($p = 0.005$).

[12]The Spearman rank correlation coefficient between stages is 0.520 ($p < 0.001$). This measure of productivity is not influenced by the treatment manipulation since subjects did not know yet during the first stage 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. Two subjects who needed 158s and 201s per table in the first stage and 46s and 58s in the main stage are not shown in Figure 1, but are of course included in all analyses. Results are unchanged if we exclude these subjects.

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

Table 1: Treatment Difference between LO- and HI-treatment in OLS and Tobit Regressions

	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**	2.065**	6.430**	6.572**	7.078**	7.927**	8.091**	8.806**
	(0.917)	(0.885)	(0.897)	(3.163)	(3.153)	(3.212)	(3.841)	(3.814)	(3.824)
Productivity		0.059***	0.062***		0.091	0.092		0.098	0.100
		(0.019)	(0.019)		(0.067)	(0.068)		(0.080)	(0.080)
1 if Female			-0.081			1.581			1.467
			(0.943)			(3.379)			(4.013)
Outside Temperature			-0.092			-0.467			-0.552
(in °C)			(0.138)			(0.496)			(0.582)
Controls for time of day	No	No	Yes	No	No	Yes	No	No	Yes
Constant	7.370***	10.607***	12.236***	31.715***	36.713***	44.615***	33.004***	38.389**	47.431***
	(0.648)	(1.206)	(3.359)	(2.237)	(4.297)	(12.031)	(2.697)	(5.143)	(14.160)
N.Obs.	120	120	120	120	120	120	120	120	120
Adjusted *or* pseudo R^2	0.03	0.09	0.09	0.03	0.03	0.01	0.00	0.01	0.01

Table 1: Treatment Difference Between LO- and HI-Treatment in OLS and Tobit Regressions. 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). 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.

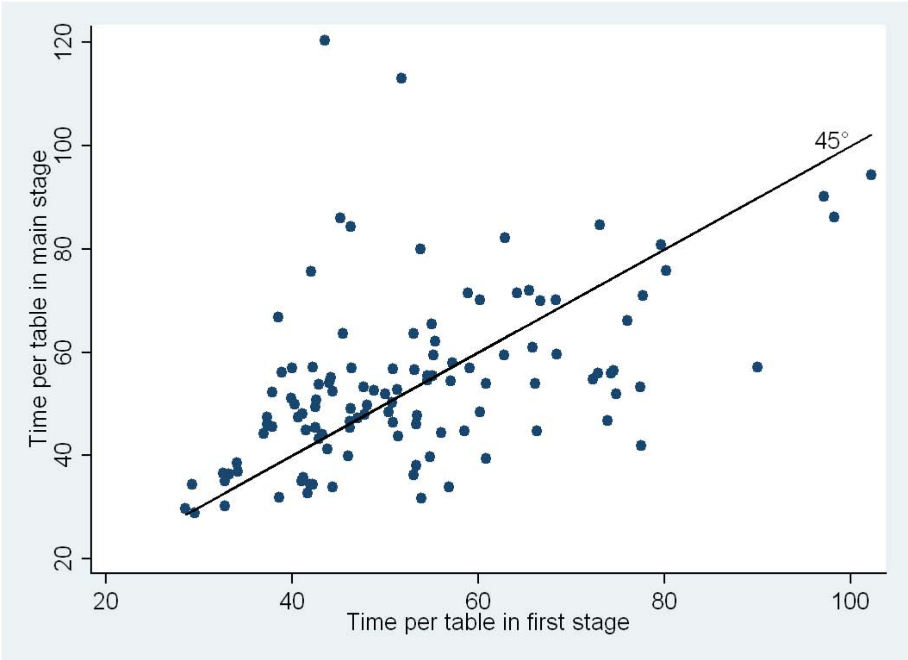


Figure 1: *Average Time per Table During First and Main Stage.*

to 9). This does not alter the results.[14]

As shown in Section 3, the model of Kőszegi & Rabin (2006) predicts that stopping decisions in the two treatments should differ in a very special way. As stated in Hypothesis 2, the model predicts a higher probability of stopping when the accumulated earnings equal the fixed payment. This prediction arises because for a whole range of marginal cost functions, stopping exactly at the fixed payment is the optimal choice. Neither the canonical model, nor the model with status quo reference 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 main treatment. The modal choice in both treatments is to stop exactly when accumulated earnings equal the fixed payment.*

Figure 2 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.

[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).

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 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 the HI-treatment many subjects stop at 7 euros (16.7 percent); in the LO-treatment 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 the OLS-regressions above 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.[16]

We conducted an additional control treatment to check whether the salience of "3 euros" and "7 euros" could have driven behavior in the two main treatments. The SAL-treatment keeps the salience of the fixed payment exactly the same as in the LO-treatment but moves the reference point away from it. If behavior in the SAL-treatment is different from the main treatment, we can conclude that the treatment difference between LO- and HI-treatment is due to reference dependence and not due to salience. Moreover, the model with the reference point in expectations makes a prediction for *how* behavior should be different: in the SAL-treatment, unlike the other treatments, there should be no special tendency to stop at the amount of the

[15]These differences are also significant in non-parametric tests: the percentages of subjects stopping at 3 euros is significantly higher in the LO-treatment (U-test, $p = 0.009$); the percentage stopping at 7 euros is higher in the HI-treatment (U-test, $p = 0.015$).

[16]Between 2 and 4 euros, 30.0 and 5.0 percent of subjects stop in the LO- and HI-treatment, 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 5 in Appendix C.

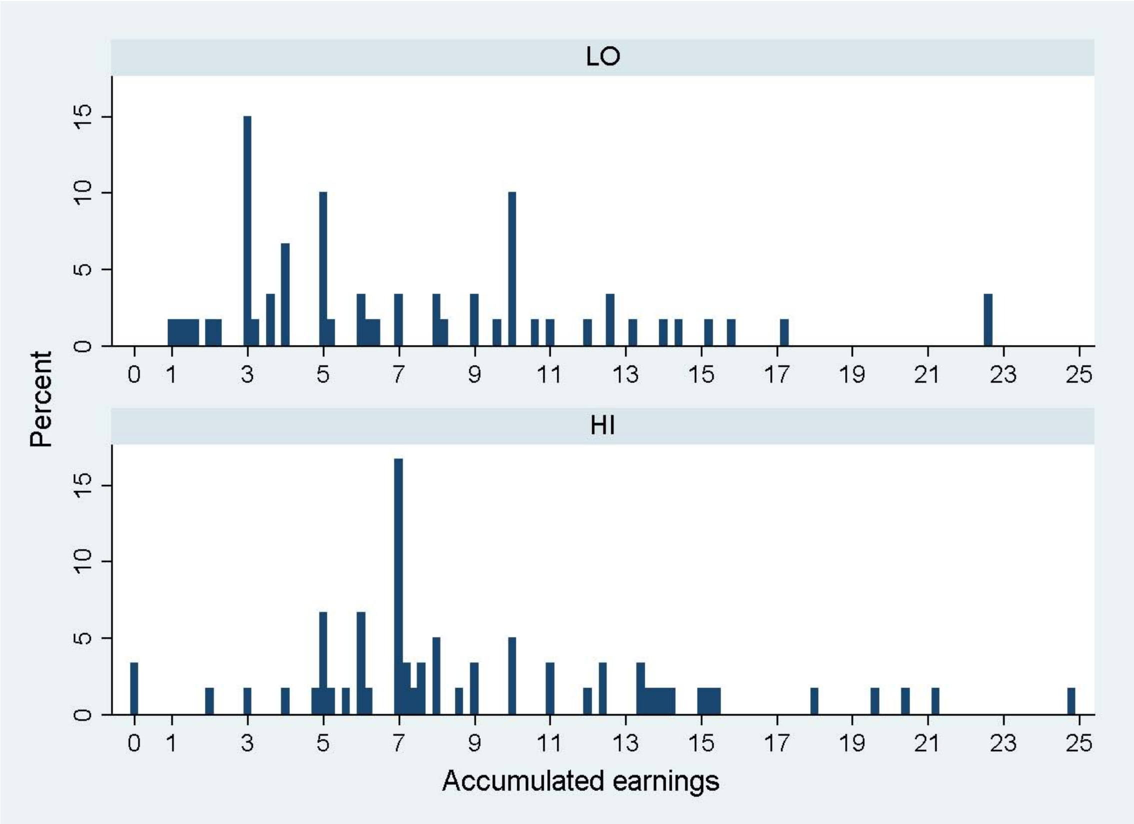


Figure 2: *Histogram of Accumulated Earnings (in Euro) at Which a Subject Stopped.*

fixed payment (Hypothesis 3). This is indeed the case.

> **Result 3:** *Subjects do not stop at the fixed payment anymore when we keep the salience of the fixed payment constant but move the reference point away from it.*

In the SAL-treatment, 3.3 percent of subjects stop at 3 euros compared to 15.0 percent in the LO-treatment. Results of multinomial logit regressions comparing the LO- and SAL-treatments are shown in Table 3. The dependent variable is again whether subjects stopped at 3 euros, at 7 euros, or somewhere else. Compared to the SAL-treatment, subjects in the LO-treatment stop more often at 3 euros but not at 7 euros.[17] This result continues to hold when we include the control variables mentioned above in the regression. Results of the control treatment therefore show that subjects in the main treatments do not stop at f because of salience.[18]

One final piece of evidence concerning why so many subjects stop exactly at the fixed payment comes from a short questionnaire administered after the two main treatments. We asked subjects to state reasons for their stopping decision. Answers were given in free form without suggestion of possible reasons. Of those subjects stopping exactly when accumulated earnings equal the fixed payment, the great majority named reasons such as a fear of "losing their earnings" when they worked for more than the fixed payment, or that they wanted to "make sure" to get the amount of the fixed payment by working at least that much. Because they indicate a desire to avoid unfavorable comparisons to what might have happened, these answers reinforce our behavioral findings suggesting that reference dependence and loss aversion drive the clustering of stopping decisions at the fixed payment.

---

[17] These differences are also significant in non-parametric tests (U-test, $p = 0.027$).

[18] A potential concern could arise because subjects in the SAL-treatment had to complete 30 tables to reach the fixed payment of 3 euros while subjects in the LO-treatment needed only 15 tables. If effort costs simply made it impossible to reach 30 tables in the SAL-treatment, this would mechanically prevent subjects from stopping exactly at 3 euros. However, 65 percent of subjects in the SAL-treatment completed at least 30 tables but only the above mentioned 3.3 percent stopped exactly at 30.

Table 2: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and HI-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**	1.609**	-2.191**	1.620**	-2.283**	1.611**
	(1.073)	(0.801)	(1.074)	(0.802)	(1.093)	(0.807)
Productivity			0.003	0.005	0.001	0.007
			(0.014)	(0.016)	(0.016)	(0.017)
1 if Female					-1.044	0.220
					(0.788)	(0.659)
Outside Temperature (in °C)					0.089	0.053
					(0.110)	(0.106)
Controls for time of day	No		No		Yes	
Constant	-1.695***	-3.199***	-1.523*	-2.946***	-3.353	-4.263*
	(0.363)	(0.721)	(0.848)	(1.121)	(2.753)	(2.556)
N.Obs.	120		120		120	
Pseudo R^2	0.09		0.09		0.12	

Table 2: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and HI-treatment). Notes: The dependent variable takes three outcomes: "stop at 3 euros", "stop at 7 euros", and "stop elsewhere" which is the reference category. Data from the LO- and HI-treatment is 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 3: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and SAL-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 LO-treatment	1.638**	0.134	1.623**	0.101	1.837**	0.414
	(0.806)	(1.019)	(0.807)	(1.022)	(0.832)	(1.077)
Productivity			0.005	0.026	0.002	0.021
			(0.014)	(0.034)	(0.016)	(0.034)
1 if Female					-0.850	-1.276
					(0.710)	(1.195)
Outside Temperature					0.128	0.152
(in °C)					(0.117)	(0.216)
Controls for time of day	No		No		Yes	
Constant	-3.332***	-3.332***	-3.038***	-2.007	-5.894*	-5.811
	(0.720)	(0.720)	(1.041)	(1.742)	(3.102)	(5.472)
N.Obs.	120		120		120	
Pseudo R^2	0.05		0.06		0.10	

Table 3: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and SAL-treatment). Notes: The dependent variable takes three outcomes: "stop at 3 euros", "stop at 7 euros", and "stop elsewhere" which is the reference category. Data from the LO- and SAL-treatments is 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.

Hypothesis 4 predicted the relation of average effort in the SAL- and LO-treatment to depend on the behavior of subjects in the LO-treatment. Average effort in SAL should be higher if most subjects in LO stopped at earnings levels above the fixed payment; and average effort should be lower if most subjects stopped below f. As shown in Figure 2, 75 percent of subject in the LO-treatment stop above the fixed payment of 3 euros and only 10 percent stop below the fixed payment. The number of subjects for whom reference dependence reduces the marginal incentive to exert effort is therefore much larger than the number of subjects for whom the marginal incentive is increased. Since the influence of reference dependence on the optimal effort choice is removed by design in the SAL-treatment, we should expect average effort in the SAL-treatment to be higher than in the LO-treatment. This is what we find.

> **Result 4:** *Subjects in the SAL-treatment work significantly more than subjects in the LO-treatment.*

Subjects complete 12.4 tables more and work 10.8 minutes longer in the SAL-treatment compared to the LO-treatment. This amounts to a marginal effect of 34 percent compared to effort in the LO-treatment. Regression analyses show that this difference is significant. Table 4 presents OLS-estimates without and with the controls described above. For both alternative dependent variables—the number of tables counted (columns 1 to 3) or the time spent working (columns 4 to 6)—subjects in the SAL-treatment work significantly more. The effect on effort provision is actually underestimated since 30.0 percent of subjects in the SAL-treatment work the full 60 minutes and are censored compared to 13.3 percent in the LO-treatment. As a consequence, the coefficient of the treatment dummy is slightly larger in Tobit estimates that take this censoring into account (columns 7 to 9).[19] The fact that subjects work more in the SAL-treatment is in line with reference dependence holding back effort in the main treatments. Once we eliminate the (low) reference

[19]The treatment differences are also significant in non-parametric tests (tables completed: U-test, $p = 0.005$; two-sample Kolmogorov-Smirnov test, $p = 0.018$. Time spent working: U-test, $p < 0.001$; Kolmogorov-Smirnov test, $p = 0.003$).

point, subjects increase effort provision drastically, even beyond the level found in the HI-treatment.

5 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. A canonical model with separable utility and models with the status-quo as the reference point fail to explain the treatment differences, whereas the results are as predicted by models with the reference point in expectations. An additional treatment ruled out an alternative explanation based on salience. 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, and have contributed to a lively debate regarding 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. Farber (2005) raised the important point 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 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 4: Treatment Difference between LO- and SAL-treatment in OLS and Tobit Regressions

	OLS: Number of tables completed			OLS: Time spent working (in min.)			Tobit: Time spent working (in min.)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1 if SAL-treatment	12.417**	13.252***	15.918***	10.804***	11.148***	13.004***	13.288***	13.586***	15.764**
	(4.781)	(4.646)	(4.933)	(3.170)	(3.148)	(3.337)	(3.957)	(3.915)	(4.069)
Productivity		0.222***	0.238***		0.092*	0.105*		0.090	0.109*
		(0.076)	(0.079)		(0.052)	(0.053)		(0.064)	(0.064)
1 if Female			-1.447			0.105			0.320
			(4.809)			(3.253)			(3.955)
Outside Temperature			-1.495			-1.042			-1.240
(in °C)			(0.938)			(0.634)			(0.762)
Controls for time of day	No	No	Yes	No	No	Yes	No	No	Yes
Constant	37.050***	49.318***	89.056***	31.715***	36.769***	63.986***	33.048***	37.967***	71.048**
	(3.381)	(5.343)	(22.609)	(2.241)	(3.620)	(15.293)	(2.759)	(4.455)	(18.431)
N.Obs.	120	120	120	120	120	120	120	120	120
Adjusted *or* pseudo R^2	0.05	0.10	0.11	0.08	0.10	0.11	0.01	0.01	0.02

Table 4: Treatment Difference Between LO- and SAL-Treatment in OLS and Tobit Regressions. Notes: The dependent variable is the number of tables completed 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). 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 are shown for OLS; pseudo R^2 for Tobit. 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.

6 References

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.

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.

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. (2007), 'The existence and position of daily income reference points: Implications for daily labor supply', *mimeo* .

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. & 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.

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

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', *mimeo* .

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* .

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.

Mehta, J., Starmer, C. & Sugden, R. (1994), 'The nature of salience – an experimental investigation in pure coordination games', *American Economic Review* **84**, 658–673.

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

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.

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 only difference in the HI-treatment is that every occurrence of "3 euros" is replaced by "7 euros". In the SAL-treatment, "3 euros" is replaced by "acquired earnings plus 3 euros" where applicable and the piece rate is set to 10 cents.

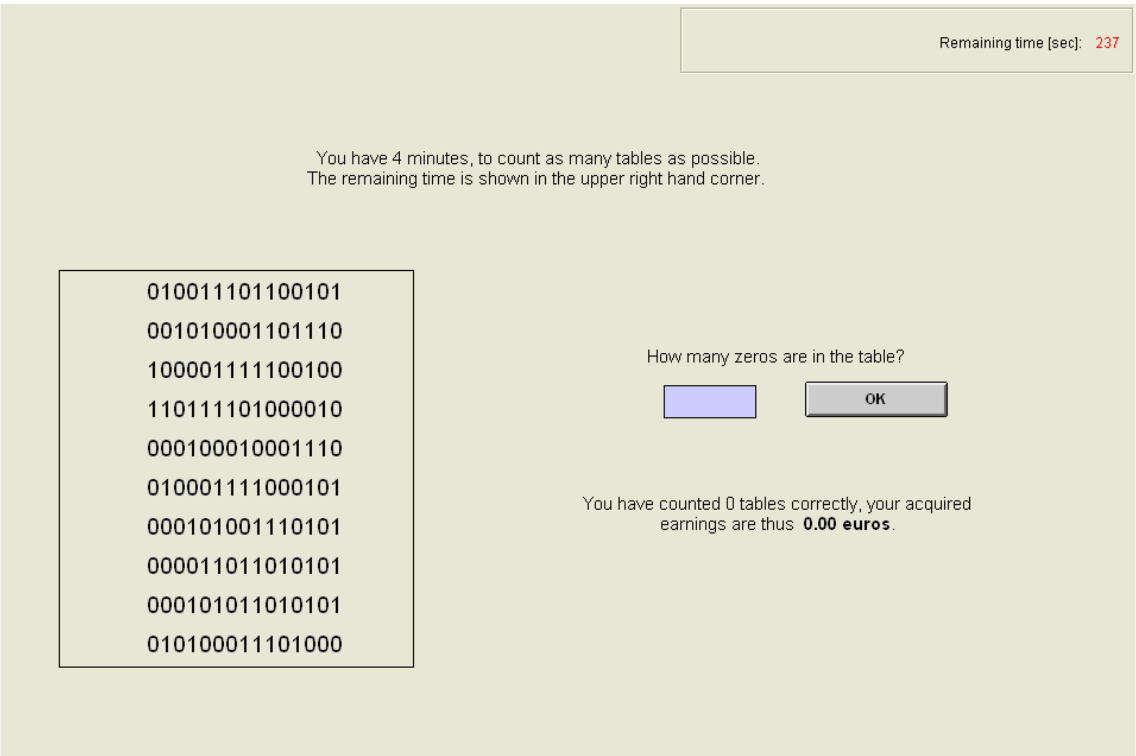
The experiment consists of two parts. Please start by reading the explanations for the first part carefully. You will receive the instructions for the second part of the experiment after the first part is finished.

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.

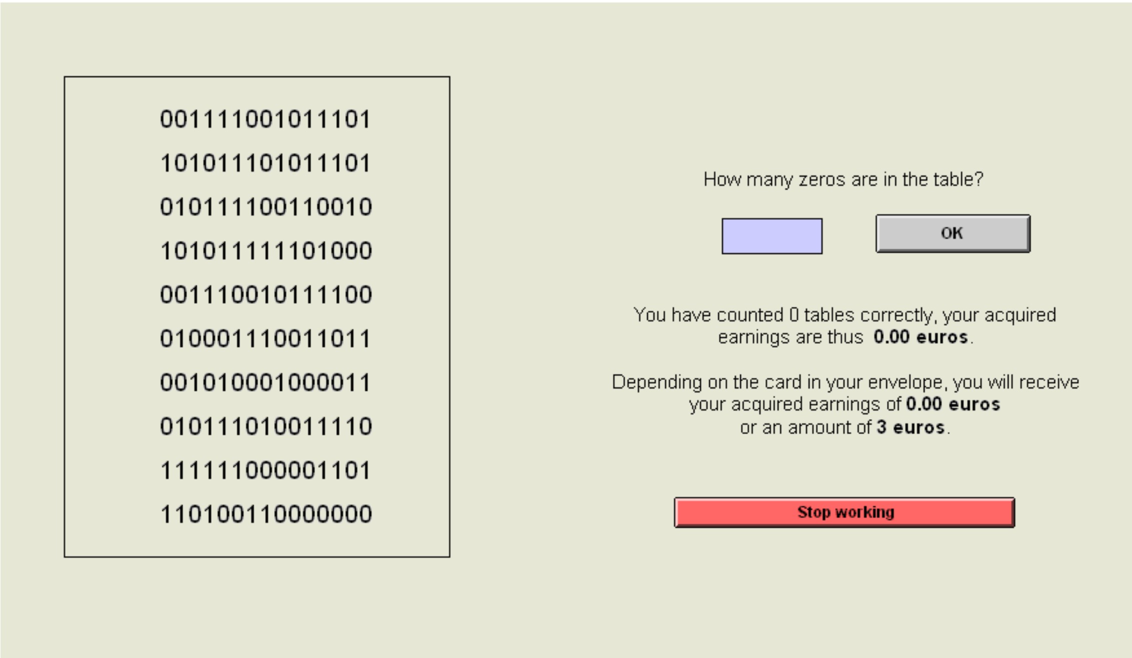
What are your acquired earnings? _____ euros

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?

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:



New rules are now in effect, which did not apply in the first part:

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

- **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.

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.

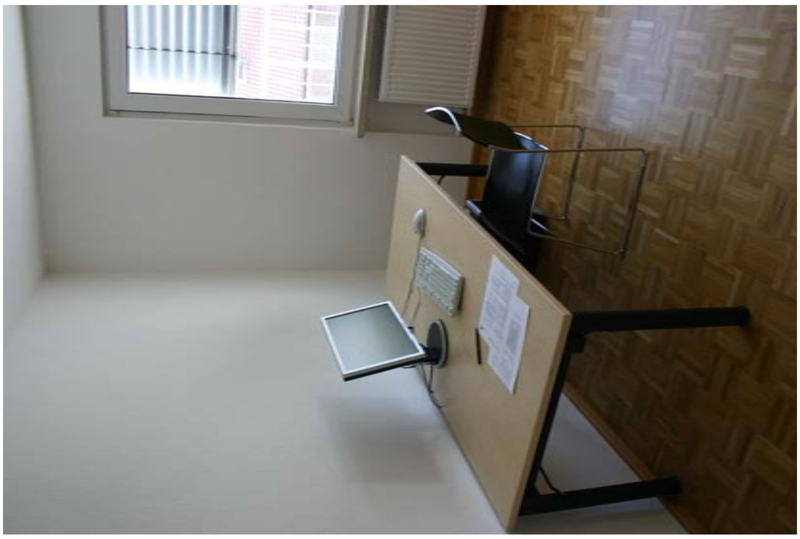
- What are your acquired earnings? _____euros
- How much money do you get if the envelope chosen by you contains the card "acquired earnings"? _____euros
- How much money do you get if the envelope chosen by you contains the card "3 euros"? _____euros

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.

B Photos of Experimental Rooms



C Additional Regression Tables

Table 5: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and HI-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***	1.056**	-1.807***	1.062**	-1.877***	1.058**
	(0.669)	(0.477)	(0.671)	(0.478)	(0.681)	(0.490)
Productivity			-0.007	0.002	-0.009	0.002
			(0.009)	(0.011)	(0.010)	(0.011)
1 if Female					-0.118	0.681
					(0.568)	(0.486)
Outside Temperature					0.051	0.018
(in °C)					(0.084)	(0.073)
Controls for time of day	No		No		Yes	
Constant	-0.636**	-1.447***	-1.033*	-1.354*	-2.184	-1.898
	(0.291)	(0.393)	(0.606)	(0.698)	(2.079)	(1.774)
N.Obs.	120		120		120	
Pseudo R^2	0.08		0.09		0.11	

Table 5: Clustering of Stopping Decisions in Multinomial Logit Regressions (LO- and HI-treatment). Notes: The dependent variable takes three outcomes: "stop between 2 and 4 euros", "stop between 6 and 8 euros", and "stop elsewhere" which is the reference category. Data from the LO- and HI-treatment is 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.