Does Reciprocity Persist Over Time?

Nickolas Gagnon* and Charles N. Noussair**

February 22, 2017

Abstract

We report the results from three distinct experiments, conducted in the Netherlands and in the United States, which extend the Gift Exchange paradigm (Fehr et al., 1993; Fehr et al., 1997) for the study of worker-employer relationships. We focus on the effect of long time delays between the time at which workers learn their wage and when they choose their effort level, on the relationship between wage and effort. We compare effort choices made on the same day workers learn their wage, with those made several weeks afterward. Our two effort choices are therefore "cold" choices. While the average effort chosen is similar under the two time lags, a positive and significant relationship between wage and effort appears consistently only in the short-term. In the long-term, the relationship is weaker and less consistent. The difference between the two time horizons is driven by the greater effort provision of workers receiving low wages. We also find more extensive reciprocal behavior by workers who receive a wage below their self-reported fair wage than those who receive more, and that this asymmetry in reciprocal behavior around the fair wage weakens after several weeks. Using a new technology that tracks facial expressions called Noldus™ FaceReader, we find that the emotion of anger is associated with reciprocal responses in the short-term, but this association is weaker in the long-term.

1. Introduction

Experimental research on the behavior of labor markets in which worker effort is not contractible, beginning with Fehr, Kirchsteiger, and Riedl (1993), has yielded a number of important insights. Even in one-shot interactions, the average wages paid to workers considerably exceed the competitive market-

* Department of Economics, Maastricht University, The Netherlands (n.gagnon@maastrichtuniversity.nl).
** Department of Economics, University of Arizona, United States (cnoussair@email.arizona.edu). The authors thank the CentER for Economic Research at Tilburg University, the Economic Science Laboratory at the University of Arizona, and the Graduate School of Business and Economics at Maastricht University, for financial support. We thank Bohan Ye for research assistance and Richard Kiser for technical assistance. We thank Kristof Bosmans, Jordi Brandts, Fortuna Casoria, Martin Dufwenberg, Peter Kuhn, Marcello Negrini, Ryan Oprea, Philippe Tobler, Agnieszka Tymula, Marie-Claire Villeval, and Matthias Wibral as well as participants at seminars at the University of Arizona, and at the ASFEE 2016 (Cergy), MBEES 2016 (Maastricht), TIBER 2016 (Tilburg), European ESA 2016 (Bergen) and North American ESA 2016 (Tucson) conferences, for helpful comments. We especially thank Gary Charness for numerous detailed comments.

clearing level. In addition, worker effort tends to be greater than the minimum effort possible, despite the fact that exerting effort is costly to the worker. A robust relationship of greater effort reciprocating for higher wages is observed. These findings have spawned a large literature (e.g. Fehr, Kirchsteiger, and Riedl, 1996; Fehr, Gächter, and Kirchsteiger, 1997; Fehr, Kirchler, Weichbold, Gächter, 1998; Fehr, Kirchsteiger, and Riedl, 1998; Fehr and Falk, 1999; Gächter and Falk, 2002; Hannan, Kagel, and Moser, 2002; Brandts and Charness, 2004; Brown, Falk, and Fehr, 2004; Charness, 2004). See Fehr, Goette, and Zehnder (2009), Charness and Kuhn (2011), Casoria and Riedl (2013), and Cooper and Kagel (2016) for surveys.

Although worker-employer relationships may be short-term or long-term in nature, the laboratory studies cited above consider short-term behavior, in the sense that all wage and effort decisions are made within one laboratory session, which is typically under two hours in duration. A number of recent studies report field experiments focused on whether the reciprocal relationship between wage and effort is durable over longer time horizons. The results with regard to the persistence of the relationship have been mixed. On one hand, Gneezy and List (2006) find that reciprocity decreases in the span of a few hours after an increase in workers' hourly wage. Bellemare and Shearer (2009) find reciprocity lasting one day after a one-off bonus payment. On the other hand, Kube, Maréchal, and Puppe (2012) observe higher effort for several hours after a one-time non-monetary gift. Gilchrist, Luca, and Malhotra (2016) observe a similar effect for the few hours following an unexpected wage increase. Cohn, Fehr, and Goette (2014) report higher productivity following an increase in hourly wage, for the entire time horizon of their study, which was several days. Hossain and List (2012) find an increase in productivity lasting for four weeks after a one-time monetary payment. Kube, Maréchal, and Puppe (2013) find no reciprocal response to a wage increase, but rather observe a sharp decrease in productivity, over a period of a few hours following a wage decrease.[1]

The mixed results obtained in these studies might be due to a number of factors, since the studies differ from each other in many aspects. Effort decisions are typically repeated, and thus some workers might exhibit a decline in effort because they feel like they have already reciprocated for any prior employer wage decisions. Worker behavior might depend on the context (List 2006; Hennig-Schmidt, Sadrieh, and

[1] In most of these studies, it is unclear whether the observed wage-effort relationship consists of positive or negative reciprocity. This is because reference wages, which are needed to classify reciprocity as positive or negative, are not elicited.

Rockenbach, 2010; Kube et al. 2012, 2013; Gilchrist et al. 2016), the perception of the fairness of the baseline wage (Fehr et al., 2009; Cohn et al., 2014), prior history (Bellemare and Shearer, 2009), anticipated future interaction between worker and employer (Hossain and List, 2012), or on cultural factors (Lensberg and van der Heijden, 1998; Henrich, Bowles, Camerer, Fehr, Gintis, and McElreath, 2001; Charness, 2004; Henrich, Heine, and Norenzayan, 2010; Hossain and List, 2012).

In the experiments reported here, we consider whether reciprocal behavior is long-term in nature, but we do so in a very simple one-shot design. We compare decisions of workers choosing their effort within three hours of learning their wage, with those of workers making their choice a few weeks after receiving their wage information. In the "hot-cold" nomenclature, both of our effort decisions are "cold" decisions.[2] There is exactly one decision made by each worker. This allows us to clearly identify and measure any and all reciprocal behavior that the worker exhibits. The experiment is decontextualized to the extent that is typical in laboratory experiments. Interactions are anonymous so that prior or future interactions between participants are not relevant to their decisions.

We conduct three separate experiments where we extend the laboratory Gift Exchange paradigm for the study of employer-workers relationships to long time horizons. The experiments differ from each other in their parametric structure and some other details, but all allow scope for a reciprocal worker-employer relationship to emerge. We consider whether workers choose effort levels to reciprocate employers' wage decisions when wages were set one month previously. In our experiments, subjects are divided into groups of three, one employer and two workers.[3] The employer chooses one wage level that applies to both workers. The first worker then chooses an effort level after leaving the laboratory, but at most three hours afterward. The second worker submits the effort decision four weeks after the experimental session. At the time their decision is due, workers choosing four weeks later receive a reminder informing them of the wage they received and the payoff structure in effect. Two of the experiments were

[2] "Hot" decisions are made instantaneously, whereas "cold" decisions allow for more reflection. See, for example, the articles by Brandts and Charness (2000) and Loewenstein (2005).

[3] Maximiano, Sloof, and Sonnemans (2007) show that the reciprocity found in Gift Exchange laboratory experiments is also observed in the case where multiple workers are matched with one employer who offers the same wage to all workers. Other studies considering the case of multiple workers and one employer in Gift Exchange settings include Guth, Konigstein, Kovács, and Zala-Mezõ (2001); Charness and Kuhn (2007); Abeler, Altmann, Kube, and Wibral (2010); Gächter and Thöni (2010); Cohn, Fehr, Herrmann, and Schneider (2014); and Charness, Cobo-Reyes, Lacomba, Lagos, Perez (2016). Most of these studies investigate social comparison effects when workers receive different wages. In our experiments, we circumvent these effects by constraining employers to offer the same wage to their two workers.

conducted in the Netherlands and one in the United States. We also test whether we observe both negative and positive reciprocity and whether one type of reciprocity dissipates more than the other with time. To distinguish negative from positive reciprocity, we ask workers what they consider a fair wage and we relate their effort to the difference between this fair wage and their actual wage. Wages lower (higher) than the fair level are assumed to lie in a negative (positive) domain for the worker, with the potential to trigger negative (positive) reciprocity.[4]

We structure our three experiments to allow application of the Partial Gift Exchange version of the Efficiency Wage hypothesis of Akerlof (1982) to the data. Our third experiment also allows a test of a key part of a later refinement, the Fair Wage-Effort hypothesis of Akerlof and Yellen (1988, 1990). The two models describe the worker-employer relationship in an environment in which effort is not contractible. In the first model, Akerlof (1982) assumes that higher wages offered by the employer are reciprocated with greater effort on the part of workers. The model predicts that employers take advantage of reciprocal effort and pay high wages, and workers reciprocate with extra effort. In the second model, Akerlof and Yellen (1990) assume that workers form an idea of what the fair wage is, and that they exhibit an asymmetric response to the wages around this fair wage. Wages below those that workers perceive as fair are punished with low effort, but those in excess of fair levels are not reciprocated in a positive manner. Employers therefore offer a fair wage to prevent shirking. They do not, however, offer a wage above the fair wage as it would not result in additional effort. Cohn et al. (2014) recently obtained support for this second model in a field experiment, finding that only the workers who felt that they were paid unfairly increased their productivity after a wage increase. Akerlof and Yellen (1990) highlight that emotions play a direct role in reciprocal behavior. They propose that anger is the psychological mechanism behind the asymmetric response to wages around the fair wage. Wages below those considered as fair trigger a reaction of anger on the part of workers, who then respond with low effort.

In one of our three experiments, we study whether worker anger is in fact negatively associated with the wage offer received, as well as with the effort subsequently exerted. Such a relationship would be consistent with observations in other contexts, where anger is correlated with negative reciprocity in immediate behavior (Bosman and van Winden, 2002; Offerman, 2002; Sanfey, Rilling, Aronson, Nystrom,

[4] Several authors have used fairness concerns to differentiate positive and negative reciprocity. In addition to Akerlof and Yellen (1988, 1990), which we discuss below, we note the work of Rabin (1993) and Dufwenberg and Kirchsteiger (2004), where negative deviations from fair or equitable behavior that are perceived as intentional generate negative reciprocity.

and Cohen 2003; Xiao and Houser, 2005; Reuben and van Winden, 2008; Hopfensitz and Reuben, 2009; Bolle, Tan, and Zizzo (2014); van Leeuwen, Noussair, Offerman, Suetens, van Veelen, and van de Ven, 2016). Battigalli, Dufwenberg, and Smith (2015) also model anger and its consequences for retaliation, using the tools of psychological game theory. They model the possibility that anger subsides over time and the desire for retaliation wanes. To measure and track anger in real time, we use face reading technology that monitors the facial expressions of participants.

We obtain six principal results in this study. First, in each of our three experiments, effort is not significantly different between workers choosing an effort several hours or several weeks after learning their wage. Second, there is a significantly positive effect of wage on effort in the short-term (hours) for all experiments, but a significant positive effect in the long-term (weeks) in only one of the three experiments. Specifically, effort in response to very low wages is lower in the short- than in the long-term. Third, negative reciprocity is more prominent than positive reciprocity, corroborating a key assumption of the Fair Wage-Effort hypothesis. Fourth, the asymmetry in reciprocity around fair wages weakens after several weeks. Fifth, the wage predictions of employers are consistent with anticipation of the short-term reciprocal wage-effort relationship when making their wage decisions. Sixth, the change in a worker's anger at the moment he learns his wage correlates with subsequent effort decisions taken within three hours, but not those taken weeks later. Indeed, those workers who know that they will receive their wage and make their effort decisions a month later, become less angry when offered a low wage.

Our results lead to five conclusions that relate to prior field experiments on Gift Exchange. First, reciprocity in worker-employer interaction seems to decrease over time, even in the laboratory, though over three hours at least, the reciprocal relationship remains strong. Second, small changes in details of the experimental design or cultural differences might affect behavior substantially. Third, we are generally in agreement with Cohn et al. (2014), who find that the effort of workers who feel that they are paid less than the wage they view as fair is more responsive to their wage. Fourth, the stronger response of workers paid below their fair wage to a wage change is less pronounced after several weeks. Fifth, anger in response to learning one's wage correlates with subsequent effort, though only in the short-term, indicating that anger is a relevant variable for understanding reciprocal behavior.

The rest of this paper is organized as follows: section 2 describes the experimental design, section 3 presents the hypotheses, section 4 reports the results, and section 5 offers our conclusions.

2. Experimental Design

In this section we describe the structure of each of the three experiments, called Tilburg-L (Tilburg Low Efficiency), Tilburg-H (Tilburg High Efficiency), and Tucson-H (Tucson High Efficiency). Section 2.1 lists the procedures that were identical in the three experiments. Section 2.2 then describes the aspects that differed among the three. It explains how the last experiment conducted, Tucson-H, allows testing of some additional hypotheses beyond those evaluated in the first two experiments.

2.1. Procedures Common to all Experiments

Participants are assigned randomly to one of three roles: 1) *Employer*, 2) *Immediate worker*, or 3) *Delayed worker*. One third of participants are assigned to each role. The roles are private information. Groups are formed, each consisting of one employer and two workers. We describe the two workers as immediate or delayed, based on when they make their effort decision. The participants are informed that the groups are randomly constituted and anonymous, in the sense that participants never learn who else was in their group. It is emphasized to them that they will play the game exactly once.

The game has two stages. In the first stage, the employer, endowed with wealth a, decides on a wage w to pay to his/her two workers. The wage must be equal for the two workers. The wage is costly to the employer and benefits the workers. In the second stage, each worker, with initial endowment d, observes his wage and chooses an effort level e. The worker's effort is costly to himself, with marginal cost c, and benefits the employer by the productivity parameter b.

The immediate worker must submit an effort level, e_I within three hours after the end of the laboratory session. The delayed worker chooses an effort level e_D within a five-day interval beginning four weeks after the session.[5] The workers, therefore, do not choose their effort in the laboratory. They send their choice to an email address provided to them, from a location of their choosing. Participants are informed

[5] The time intervals in experiment Tilburg-L are slightly different than in the other two experiments. Immediate workers were instructed to send their effort choice within two hours after the session, and delayed workers during a 24-hour window that began one month after the session. This was not enough time for both types of workers, usually because some students check their university mailbox infrequently and therefore missed the reminder during the 24-hour window one month after the session. We accepted effort choices from a few immediate workers after the two-hour deadline. We also informed delayed workers by email that they would have additional days to choose their effort and accepted choices up to eight days after the original deadline.

during the session that the delayed workers would receive an email reminder, containing the wage they have been awarded and a copy of the instructions.[6] [7] Participants are also informed that if a worker does not send an email with her effort choice, she would not be paid anything beyond the show-up fee and the employer would be paid back the wage offered to that worker.[8]

The earnings of workers are sent to them on the day after the receipt of their effort choice. The earnings of the employer are sent on the day after the receipt of the effort choice of the delayed worker she is matched with.

The payoffs to the three types of participants were given by:[9]

Employer**:** $a - 2 \times w + b \times (e_I + e_D)$

Immediate worker: $d + w - c \times e_I$

Delayed worker: $d + w - c \times e_D$

The variables e_I and e_D denote the effort levels chosen by the immediate and delayed workers, respectively. The ranges of possible wages, effort levels, effort costs and output differed among the three experiments. The values are given in Table 1.

Wages and effort can be chosen in increments of 10 cents and 0.1 units, respectively. Note that the ratio *b/c*, denoting the benefit of effort to the employer over the cost of effort to the worker, is constant within each experiment (Gift Exchange experiments often feature a diminishing *b/c* ratio). We use a constant ratio for simplicity, which is especially desirable since participants only play the game once. Sessions generally had between 9 and 15 participants, and a few had 6 participants. Demographic statistics for the participants are presented in Table 2.

---

[6] The instructions for Tucson-H, which are similar to those for Tilburg-L and Tilburg-H, are in Appendix C.
[7] The email reminders for each experiment are in Appendix D.
[8] Paying back the wage to employers when workers do not choose an effort does not change the fact that the least costly action of workers is choosing an effort of zero, and the action of workers that earns the employer the most is choosing maximum effort.
[9] The payoff structure was presented to subjects in terms of these formulas, with the actual values in effect, rather than the variables *a, b, c, d*, indicated. See Charness, Fréchette, and Kagel (2004) for a discussion of the effects of different formats of presentation.

Table 1 - Summary of Parameters

Parameters & Range of Variables	Tilburg-L	Tilburg-H	Tucson-H
a	12 EUR	8 EUR	11 USD
b	2.5	2.5	2.5
c	1	0.25	0.25
d	4 EUR	5 EUR	4 USD
e	[0, w]	[0, 2.4]	[0, 4]
w	[0 EUR, 4 EUR]	[0 EUR, 2 EUR]	[0 USD, 4 USD]
Show-up Fee	0 EUR	0 EUR	8 USD
Average Earnings/Employer	10 EUR	8 EUR	22 USD
Average Earnings/Imm. Worker	6 EUR	6 EUR	14 USD
Average Earnings/Del. Worker	6 EUR	6 EUR	14 USD
Average Duration of Session	40 Min.	40 Min.	40 Min.
Payment Method	Bank Transfer[10]	Bank Transfer	Cash and Mail
Time Period Conducted	March-May 2015	April-June, November-December 2015	January-March 2016
Location	CentER Lab (Tilburg, The Netherlands)	CentER Lab (Tilburg, The Netherlands)	Economic Science Lab (Tucson, United States)
Number of Participants	138	99	194[11]
Attrition/Imm. Workers[12]	0%	3%	7.7%
Attrition/Del. Workers	9%	9%	12.5%
Net Number of Participants	134	95	181

A session proceeds in the following manner. Participants arrive at the laboratory and are seated individually at a computer. They are given a written copy of the instructions, which the experimenter reads aloud. Participants are forbidden from communicating with others. Before making their choices, they must complete two practice exercises to confirm their understanding. They are encouraged to ask questions and their individual answers are verified by the experimenter. Help is provided if needed, as each participant is required to fill in the correct answers to proceed further.[13]

[10] Bank transfers are the most common method of payment in the Netherlands, even for small payments.
[11] One subject participated twice. We removed the observation from the second participation.
[12] Attrition is not correlated with low wage offers. In Tucson-H, the mean wages of immediate and delayed workers who did not respond were 2.30 USD and 2.60 USD, close to the overall sample mean of 2.62 USD. In Tilburg-L, the four delayed workers who did not respond had a mean wage of 1.58 EUR, slightly below the sample average of 2.07 EUR. In Tilburg-H, one immediate worker with a wage of 0.70 EUR did not respond, and three delayed workers with a mean wage of 0.90 EUR did not respond. This is, again, close to the mean sample wage of 1.22 EUR.
[13] The game was programmed using the z-Tree software (Fischbacher, 2007).

Table 2 – Demographic Profile of Participants

Demographics	Tilburg-L	Tilburg-H	Tucson-H
Age	22.9	23.3	20.8
Male	50%	53%	52%
At least 1 Friend in Session	38%	29%	13%
At least 2 Friends in Session	17%	2%	4%
At least 3 Friends in Session[14]	6%	0%	1%
Business/Economics Major	64%	67%	71%
Science Major	2%	1%	20%
Social Science Major	31%	32%	8%
Other Majors	4%	1%	5%
American	.	.	78%
Asian	28%	15%	19%
European	64%	72%	.
Other	9%	15%	3%

Majors and countries/continents do not sum up to one because of double majors and dual nationals.

2.2. Procedures Specific to Each Experiment

2.2.1 The Tilburg-L and Tilburg-H Experiments

The procedures for the Tilburg-L experiment are described fully in section 2.1. The Tilburg-H experiment differs from Tilburg-L in several respects. As in standard Gift Exchange experiments, the effort choice of workers is not limited by the wage offered. The parameters in effect in Tilburg-H also ensure greater productivity of effort, by both reducing the cost and increasing the output of effort. This is captured in the ratio *b/c*. A relatively high ratio tends to reduce the censoring of the effort data at the lower bound (Kessler, 2013).

There are two other important differences between Tilburg-L and Tilburg-H. In Tilburg-H, employers are asked to make predictions about the choices of both the immediate and the delayed workers that they are matched with. One of the two predictions is selected to count toward their payment. They earn 1 EUR if the prediction selected to count is within 0.2 units of the actual effort chosen. At the same time that he chooses a wage, the employer also requests a non-binding effort level from each worker. The effort level

[14] In Tilburg-L, four participants claimed to have four friends in a session. In Tucson-H, a participant claimed to have many friends in a session in which all other participants stated they had none.

requested is constrained to be the same for the two workers. The effort request is included in the reminder to delayed workers to ensure that they can recall it.

2.2.2 The Tucson-H Experiment

This experiment differs from Tilburg-H in several respects. There are a number of differences in the parameters between the two experiments, as indicated in Table 1. In Tuscon-H, participants receive a show-up fee of 8 USD at the time of the session on top of their earnings from the game. However, the key ratio *b/c* remains the same as in Tilburg-H, at a level of 10. Employers also predict the effort levels of the two workers that they are grouped with, but they receive 1 USD if their selected prediction of worker effort is within 0.4 of the actual effort. Furthermore, in Tucson-H, employers request a non-binding effort level from each worker at the time that they announce the wage they are paying. This request is constrained to be the same for each worker. In addition, all workers are asked to state what they think would be a fair wage. They do so before learning their actual wage. It is emphasized to them that this fair wage would not be revealed to any other participant and therefore could not affect the actual wage they receive.

We videotape participants for the entire session, with their consent. The videotapes are then analyzed with Noldus FaceReader software, which tracks facial expressions and analyzes the emotions they express. The software tracks facial movements using the Facial Action Coding System, which associates specific muscle movements to the six basic universal emotions catalogued by Paul Ekman and his colleagues (e.g. Ekman and Friesen, 1977; Ekman et al., 1987; and Ekman and Friesen, 2003). The emotions are happiness, fear, anger, disgust, surprise, and sadness. FaceReader also measures how closely a facial expression conforms to a neutral state and generates an overall measure of emotional valence. FaceReader has been employed in a number of experimental economic studies focusing on emotions (e.g. Nguyen and Noussair, 2014; Breaban and Noussair, 2014; Habětínová and Noussair, 2015; Van Leeuwen et al., 2016), and has also been used in marketing (Teixeira, Wedel, and Pieters, 2012; Lewinski, Fransen, and Tan, 2014) and psychology research (Chentsova-Dutton and Tsai, 2010; Gadea, Aliño, Espert, and Salvador, 2015).

In the Tucson-H experiment, participants are paid in cash. The participation fee is paid at the end of the session. Earnings from the Gift Exchange game are sent by mail in cash using an envelope that participants are given at the beginning of the session, on which they write their name and mailing address.

3. Hypotheses

The first two hypotheses apply to each of the three experiments, and concern the effect of a long vs. a short time delay on effort choices. They originate from the mixed results of prior experimental work, and compare the behavior of immediate and delayed workers. Hypotheses 3 and 4 apply exclusively to Tucson-H. They emerged after the first two experiments and guided the design of Tucson-H. They posit asymmetries in worker behavior between cases in which the wage is perceived as fair and as unfair. Hypothesis 3 is an assumption in the model of Akerlof and Yellen (1988, 1990) that has also been evaluated by Cohn et al. (2014). This assumption states that a reciprocal wage-effort relationship only holds for wages below a level considered fair. Hypothesis 4 asserts that this asymmetry in behavior above vs. below the fair wage holds independently of the time delay.

We first consider the relationship between effort and time delay, for a given wage. It is not clear from the literature whether we should expect immediate and delayed workers to choose different effort levels and in which direction these levels might differ. The most closely related studies are Kovarik (2009) and Dreber et al. (2016), who investigate the interaction between altruism and long time-lags. They find that when dictators have to split money that is payable in the future, they keep more for themselves. This is, however, different from the choice workers face here, which is governed in part by altruism, but also by reciprocity and perhaps other considerations. Furthermore, in our experiments, workers are reacting to a past decision by another agent rather than making a decision affecting the future. In the absence of prior research on closely related questions, we propose the following null hypothesis.

Hypothesis 1: For a given wage, immediate workers exert the same effort as delayed workers ($e_I|w= e_D|w$).

While hypothesis 1 concerns the level of effort, hypothesis 2 considers whether a positive relationship between wage and effort appears with similar force in the short and long term. The Gift Exchange version of the Efficiency Wage Hypothesis (Akerlof, 1982) does not consider a potential interaction of reciprocity and time. As discussed in the introduction, field experiments on whether reciprocity is sustained over time have found diverse results. Similar mixed results are found in laboratory studies of the effect of short time delays on reciprocity. Grimm and Mengel (2011) find that a 10-minute delay reduces rejections in ultimatum games. Neo et al. (2013) observe a reduction of rejections in ultimatum games with a similar delay and no effect on reciprocal behavior in the investment game. Thus, our second hypothesis is as follows.

Hypothesis 2: Immediate and delayed workers respond with the same change in effort to a change in their wage ($de_I/dw = de_D/dw$).

Hypothesis 3 is a test of the Fair Wage-Effort Hypothesis (Akerlof and Yellen 1988, 1990). The hypothesis relies on the assumption that workers respond to a higher wage with higher effort only up to the point where the wage is considered fair. Workers do not reciprocate for an increase in wage above the fair wage with greater effort. Cohn et al. (2014) are, to our knowledge, the only authors who test whether there is an asymmetric response at the wage viewed as fair, as assumed in the Fair Wage-Effort Hypothesis. In Tucson-H, as explained in Section 2, we ask workers for a self-reported fair wage before they receive their actual wage. Because there is only one previous study on the topic, and the study differs in important aspects from ours, we maintain as our initial hypothesis that the relationship between wage and effort is the same for workers whose wage is below vs. above their belief about what constitutes a fair wage.[15]

Hypothesis 3: Immediate workers receiving a wage below their self-reported fair wage respond with the same change in effort to a change in their wage as those receiving a wage above their fair level (de_I/dw_L, $= de_I/dw_H$ for all w_L and w_H, where $w_H > w_F > w_L$ and w_F is the fair wage).

Hypothesis 4 concerns whether any asymmetry in reciprocity below vs. above the fair wage continues to exist if there is a delay in effort choices. It is possible that negative reciprocity dissipates over time whereas positive reciprocity does not. This could be the case if, for example, negative reciprocity is caused by a strong negative emotion, such as anger, that would dissipate over time. Indeed, Akerlof and Yellen (1990) present anger as the driver of the negative response of workers to getting less than what they view as a fair wage. The FaceReader data allow us to measure overall emotional valence and anger, and study their relationship to negatively reciprocal decisions. In the absence of any prior evidence to the contrary, our null hypothesis is that any asymmetry in reciprocity based on whether workers receive a wage below or above their fair wage is unaffected by a delay.

Hypothesis 4: Immediate workers and delayed workers respond with the same relationship between their effort and their wage when (a) the wage they receive exceeds their reported fair wage, and when (b) the

[15] Considering that negative reciprocity has been repeatedly found to be stronger with other reference measures than the self-reported fair wage, our hypothesis could also be one-sided.

wage they receive is less than their reported fair wage ((a) $de_I/dw = de_D/dw$ for $w > w_F$, and (b) $de_I/dw = de_D/dw$ for $w < w_F$).

4. Results

This section is organized in three parts. The first part presents an overview of the data. The second reports the tests of our hypotheses, and the third is an analysis of other patterns in our data, including the relationship between emotions and effort decisions. Our first four results, presented in section 4.2, appear in the same order as the hypotheses they evaluate. Results 5 and 6, reported in section 4.3, concern the connection between anger and low effort, as well as an analysis of employers' predictions.

4.1. Summary of the Data

In Table 3, we summarize the average values of participants' decision variables.

Table 3 – Summary of Participants' Decisions

Experiment	Tilburg-L	Tilburg-H	Tucson-H
Wage	2.067 (1.288)	1.218 (0.683)	2.617 (1.014)
Effort of Immediate Worker	0.343 (0.595)	0.538 (0.727)	1.395 (1.217)
Effort of Delayed Worker	0.343 (0.596)	0.61 (0.833)	1.704 (1.397)
Fair Wage Indicated by Imm. Worker			3.134 (0.801)
Fair Wage Indicated by Del. Worker			3.413 (0.716)
Predicted Effort of Imm. Worker		1.142 (0.684)	2.163 (0.919)
Predicted Effort of Del. Worker		1.137 (0.703)	1.958 (1.258)

The numbers presented are averages. In Tilburg-L and Tucson-H, wage and effort range from 0 to 4. In Tilburg-H, wage ranges from 0 to 2 and effort from 0 to 2.4. Standard errors are in parentheses.

The table reveals the following patterns. The average effort of immediate and delayed workers is similar, and is well in excess of the minimum of zero. Delayed workers exert more effort than immediate workers in Tilburg-H and Tucson-H, but the differences are not significant. The average wage is close to the midpoint of the range of possible wage levels. The wage workers view as fair is roughly three times as far from the minimum as the maximum possible wage. The fair wage exceeds the actual average wage. The fair wage is similar for immediate and delayed workers. Employers' predictions about workers' effort are more optimistic than the actual observed effort, and are similar for the immediate and the delayed workers.

Table 4 – Average Effort by Wage Range, for Immediate and Delayed Workers

Wage Range (EUR)	*Immediate Workers*				*Delayed Workers*			
	Average Effort	N Min. Effort	N Max. Effort	N	Average Effort	N Min. Effort	N Max. Effort	N
Tilburg L								
0-1	0.015 (0.038)	11	0	13	0.018 (0.041)	9	0	11
1.1-2	0.364 (0.607)	6	0	11	0.74 (0.648)	2	0	10
2.1-3	0.48 (0.718)	3	0	10	0.13 (0.200)	5	0	10
3.1-4	0.567 (0.702)	5	0	12	0.5 (0.834)	6	0	11
Tilburg H								
0-0.5	0.1 (0.127)	3	0	6	1.38 (1.226)	1	2	5
0.6-1	0.2 (0.187)	1	0	5	0.283 (0.382)	2	0	6
1.1-1.5	0.491 (0.592)	4	0	11	0.278 (0.559)	5	0	9
1.6-2	1.02 (0.976)	1	0	10	0.72 (0.828)	4	1	10
Tucson H								
0-1	0.167 (0.408)	5	0	6	1 (1)	3	0	7
1.1-2	0.85 (0.880)	3	0	12	1.209 (1.221)	3	1	11
2.1-3	1.443 (0.947)	0	0	23	1.871 (1.420)	5	4	21
3.1-4	2.068 (1.435)	3	3	19	2.106 (1.513)	3	4	17

The minimum and maximum possible efforts, as well as the wage range, differ between experiments. We indicate the number of participants choosing the minimum possible effort (N Min. Effort), the maximum possible effort (N Max. Effort), and the total number of participants (N). Standard errors are in parenthesis.

Table 4 shows the average effort levels for different wage ranges, for immediate and delayed workers. It also displays how many workers choose the minimum effort level of 0, how many choose the maximum effort level, and the number of workers receiving a wage in each wage range. The table reveals that for immediate workers, there is a rather strong tendency to reciprocate a higher wage with higher effort. This increasing relationship between wage and immediate effort appears in each of the three experiments. The wage-effort relationship seems to be weaker for delayed workers, especially in Tilburg-L and Tilburg-H. Whereas low wages (in the first and second quarters of the wage range) lead to very low effort levels for immediate workers, they lead to greater effort for delayed workers. This tendency is less pronounced In Tilburg-L, where the effort choices of workers are bounded above by the wage offered, so that is impossible for workers with very low wages to select high effort levels. There are also fewer instances of choosing the minimum effort level for low wages among delayed workers. In addition, there are more instances of choosing the maximum effort level.

Figure 1 – Scatter Plots of Immediate and Delayed Efforts for Each Experiment

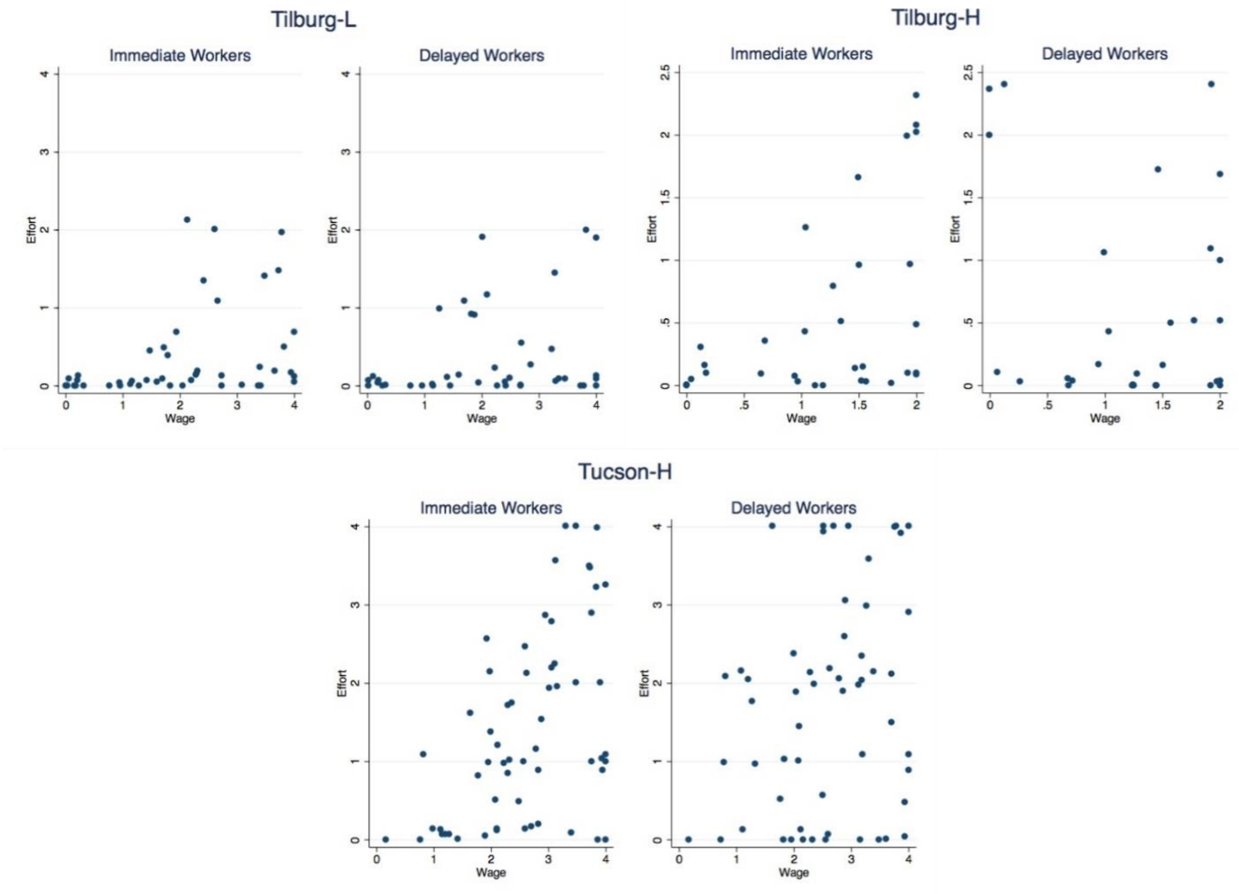

Figure 1 illustrates the relationship between the wages workers receive and the effort they choose. The two panels in the upper left contain the data from the Tilburg-L experiment, while the upper right and lower panels show the data from Tilburg-H and Tucson-H. The wage-effort relationships are graphed separately for immediate and delayed workers. One can observe the increase in effort for delayed workers with low wages, especially in Tilburg-H and Tucson-H, where efforts are not bounded above by the wage.

Thus, our three experiments replicate the two main findings of the Gift Exchange literature. (1) Both wages and effort are considerably higher than (0, 0), the subgame perfect equilibrium levels under the assumption of pure self-interest. (2) For immediate workers, higher wages are correlated with greater effort.

4.2 Evaluation of Hypotheses

The first two hypotheses concern differences in effort between immediate and delayed workers. To test these two hypotheses, we estimate the following Tobit specification.[16] The Tobit model is chosen to account for the censoring of the effort variable:[17]

$$\boldsymbol{e_j = \alpha + \gamma \times D_j + \beta_{imm} \times wage_j \times (1 - D_j) + \beta_{del} \times wage_j \times D_j + \varepsilon_j} \qquad (I)$$

In this equation, e_j denotes the effort of worker *j*, $wage_j$ denotes the wage, and D_j equals 0 for immediate workers and 1 for delayed workers. We also include a set of demographic controls consisting of the worker's age, gender, nationality and whether one has at least one friend in the session. The estimates are presented in Table 5.

To test our first hypothesis, that the time delay has no effect on effort for a given wage, we use an F-test for the joint restrictions that $\gamma = 0$ and $\beta_{imm} = \beta_{del}$ for each experiment. The F-test fails to reject the hypothesis for each of the three experiments, although the p-values are often not far from significance levels. For Tucson-H and Tilburg-H, the p-values are 0.208 and 0.092 without controls and 0.127 and 0.111 with demographic controls, respectively. For Tilburg-L, the p-values are 0.496 without controls and 0.811 with demographic controls. We can also test for differences in the average effort of workers, which are shown in Table 3. The average effort is slightly greater for immediate than delayed workers in Tilburg-H and Tucson-H (0.61 versus 0.538 and 1.704 versus 1.395). The differences are not significant when t-tests are conducted (p-values are 0.716 and 0.207).[18] The average effort is exactly the same for Tilburg-L (0.343). This yields our first result:

Result 1*: For a given wage, immediate and delayed workers do not choose significantly different effort levels.*

To test our second hypothesis, that the wage-effort relationship does not change with delay, we use an F-test of the restriction $\beta_{imm} = \beta_{del}$ for each experiment. The restriction is significant for Tilburg-H, but not for Tilburg-L or Tucson-H (p-values are 0.030, 0.277, and 0.382 without demographics and 0.039,

[16] We use robust standard errors in all regressions.
[17] In Tilburg-L, we assume for simplicity that only effort choices of 0 are censored below in the Tobit regression, even though all effort levels are subject to censoring from above if they equal the wage.
[18] Mann-Whitney tests also reveal no differences (p-value is 0.898 for Tilburg-H and 0.341 for Tucson-H).

Table 5 – Estimates of the Effect on Effort from Tobit Regression (I)

Experiment	(1) Tilburg-L	(2)	(3) Tilburg-H	(4)	(5) Tucson-H	(6)
Wage/Imm. Workers (β_{imm})	0.422***	0.325***	0.729***	0.750**	0.918****	0.933****
	(0.125)	(0.116)	(0.239)	(0.296)	(0.233)	(0.205)
Wage/Del. Workers (β_{del})	0.217	0.224	-0.396	-0.368	0.606**	0.626**
	(0.143)	(0.138)	(0.439)	(0.441)	(0.271)	(0.263)
Delay (γ)	0.499	0.124	1.453**	1.412**	1.240	1.308
	(0.425)	(0.393)	(0.702)	(0.689)	(0.916)	(0.859)
Constant	-1.050***	-4.858****	-0.598*	-2.032	-1.213**	1.433
	(0.310)	(1.082)	(0.319)	(1.405)	(0.596)	(1.540)
Age		0.181****		0.075		-0.125*
		(0.042)		(0.048)		(0.069)
Man		-0.049		-0.150		0.497
		(0.218)		(0.313)		(0.330)
Friend		-0.034		0.054		1.305***
		(0.233)		(0.395)		(0.411)
American						-0.656*
						(0.351)
European		-0.050		-0.351		
		(0.212)		(0.389)		
Pseudo-R^2	0.064	0.190	0.047	0.080	0.055	0.086
N	88	88	62	62	116	116

N denotes the sample size containing both types of workers. Workers are split almost even per type. Standard errors are in parentheses. P-values: *0.10, **0.05, ***0.01, ****0.001.

0.565, and 0.358 with demographics, respectively). We also test whether a reciprocal relationship is present at all with t-tests of $\beta_{imm} = 0$ and $\beta_{del} = 0$ separately for each experiment. The t-test soundly rejects that $\beta_{imm} = 0$ for each of the three experiments (p-values are 0.001 for Tilburg-L, 0.003 for Tilburg-H, and <0.001 for Tucson-H without demographics, and 0.006, 0.014, and <0.001 with demographics). The t-test similarly rejects that $\beta_{del} = 0$ for Tucson-H (p-value is 0.027 without demographics, and 0.019 with demographics), but fails to reject that $\beta_{del} = 0$ for Tilburg-L and Tilburg-H (p-values are 0.133 and 0.371 without demographics, and 0.109 and 0.408 with demographics, and the coefficient for Tilburg-H is negative). Note, however, that even in Tucson-H the coefficient of wage for delayed workers is only two-thirds of the coefficient for immediate workers. When estimating a Tobit regression on the Tucson-H data for immediate and delayed workers separately, we also observe that the pseudo-R^2 is much larger for immediate than for delayed workers (0.096 and 0.026 without demographics, and 0.182, and 0.041 with demographics).[19] [20]

Table 6 – Correlation between Wage and Effort, for Immediate and Delayed Workers

Wage-Effort Correlation	Tilburg L	N	Tilburg H	N	Tucson H	N
Immediate Workers	0.367 [0.004]	46	0.468 [0.012]	32	0.494 [<0.001]	60
Delayed Workers	0.248 [0.306]	42	- 0.191 [0.681]	30	0.296 [0.026]	56

We indicate the two-sided p-values from Spearman's rank correlation tests in square brackets.

We can also employ non-parametric tests to investigate differences in the wage-effort relationships. In Table 6, we show the correlations between wage offered and effort chosen for immediate and delayed workers in each experiment. We perform Spearman's rank correlation tests on each of the correlation coefficients and indicate the p-values in square brackets. We can observe that the wage-effort correlations are always positive and significant for immediate workers (p-values are all <0.02). The same is not true for delayed workers. For delayed workers, the wage-effort correlations are smaller in magnitude and not significant for Tilburg-L and Tilburg-H (even negative for Tilburg-H), and smaller but significant for Tucson-H. Another, stricter, test is to examine whether the correlations differ significantly in the short- and long-term. The difference is significant for Tilburg L only (Fisher's r to z transformation, two-sided p-values are 0.549 for Tilburg-L, 0.009 for Tilburg-H, and 0.215 for Tucson-H).

Specifically, the average effort in the first quarter of the wage range is significantly higher for delayed workers than for their immediate colleagues. The difference is significant or marginally significant for two of the three experiments, using one-sided t-tests on the average (one-sided p-values are 0.432 for Tilburg-L, 0.040 for Tilburg-H, and 0.039 for Tucson-H), or one-sided Mann-Whitney tests on the distribution (one-sided p-values are 0.429, 0.067, and 0.05). It is perhaps not surprising that the effort is not significantly higher for Tilburg-L, considering that workers' effort choices are restricted above by the low wage received. In contrast, the effort for wages in the pooled upper three quarters of the wage range do not differ significantly for delayed workers compared to immediate workers (two-sided p-values from t-tests

[19] A similar pattern of greater explanatory power of wage on the effort provision for immediate than for delayed workers is observed in Tilburg-L and Tilburg-H.

[20] Our results for Tucson-H are robust to excluding participants who have a friend in a session.

are 0.930 for Tilburg-L, 0.369 for Tilburg-H, and 0.299 for Tucson-H; two-sided p-values from Mann-Whitney tests are 0.911, 0.272, and 0.463).

The above analysis provides the basis for our second result:

Result 2*: For immediate workers, higher wages are reciprocated with greater effort. The relationship is weaker and less consistent after a one-month time delay. This difference is driven by the greater effort of delayed workers receiving low wages.*

Our two remaining hypotheses apply only to Tucson-H. They both concern the relationship between effort chosen and the gap between self-reported fair wages and actual wages. In the upper part of Table 7, we present the average efforts chosen for different ranges of fair and actual wages.[21] The upper part of the table depicts two main features. First, for wages substantially below the fair wage, higher wages lead to greater effort, especially for immediate workers. Above the range around the fair wage, higher wages do not seem to lead to much higher effort levels than wages close to the fair wage do. The precision of this second observation, however, is limited by the relatively low number of participants receiving a wage in excess of the fair wage. Second, effort levels in response to wages below fair levels are higher for delayed than for immediate workers.

Table 7 - Average Effort for Different Ranges of the Gap between Actual and Fair Wages, for Immediate and Delayed Workers

	Immediate Workers		*Delayed Workers*	
Wage - Fair Wage (EUR)	Average Effort	N	Average Effort	N
[-4; -1.6]	0.660 (0.844)	10	1.175 (1.085)	16
[-1.5; -0.6]	0.977 (0.980)	17	1.738 (1.286)	16
[-0.5; 0.5]	1.755 (1.192)	22	2.079 (1.710)	19
[0.6; 1.5]	2.322 (1.359)	9	1.860 (1.113)	5
[1.6; 4]	0.500 (0.707)	2	.	0

Correlation between (Wage - Fair Wage) and Effort

	< Fair Wage	N	≥ Fair Wage	N
Immediate Workers	0.460 [0.001]	38	-0.073 [0.645]	22
Delayed Workers	0.211 [0.263]	39	-0.057 [0.945]	17

Standard errors are in parentheses. We indicate the two-sided p-values from Spearman's rank correlation tests in square brackets.

The third hypothesis is that immediate workers would exhibit the same wage-effort relationship for wages below and above their self-reported fair wage. The lower part of Table 7 shows the correlations between

[21] We provide a more detailed table in Appendix A.

the fair wage gap (wage - fair wage) and effort, for wages both below and above the fair wages. Wages exactly equal to the fair wage are included with the wages above.[22] We start by focusing only on immediate workers. The correlation between effort and the fair wage gap is always positive and significant for immediate workers paid below their fair wage (correlation is 0.460; p-value from Spearman's rank correlation test is 0.001). The correlation between effort and the fair wage gap is near zero and insignificant for wages above or equal to fair levels (correlation is -0.073; p-value is 0.645). As a more stringent test, we find that the correlations below and above the fair wage are significantly different (p-value from Fisher's r to z transformation is 0.0455).

To take into account both the effect of the wage and the fair wage gap on effort, we present a more detailed parametric analysis of the immediate workers' effort, for which we estimate the following Tobit regression:

$$\boldsymbol{e_j = \alpha + \delta_{imm-above} \times z_j + \beta_{imm-above} \times \left(wage_j - fairwage_j\right) \times z_j + \beta_{imm-below} \times \left(wage_j - fairwage_j\right) \times \left(1 - z_j\right) + \lambda \times wage_j + \varepsilon_j} \qquad (\boldsymbol{II})$$

Here, *(wage*$_j$ *– fairwage*$_j$*)* denotes the difference between a worker's actual wage and her self-reported fair wage. The dummy variable z_j takes value 1 if the wage is equal to or above the fair wage and 0 if the wage is below the fair level submitted by worker *j*. The estimates are presented in columns 1 and 2 of Table 8.

The negative sign of $\beta_{imm-above}$ indicates that for immediate workers, the wage-effort relationship is less steep for wages above the fair level than below. To test for equality of the wage-effort relationship above vs. below the fair wage, we use an F-test of the restriction that $\beta_{imm-above} = \beta_{imm-below}$. The F-test does not reject that the two coefficients are equal, although the one-sided test without demographics is close to a significant level (the p-value is 0.118 without demographics and 0.208 with demographics). A less stringent test is to examine whether wage-effort reciprocity is significantly different from zero, below and above the fair wage separately. To do this, we first test whether there is reciprocity when there is a

[22] According to Akerlof and Yellen (1990), paying fair wages should elicit the same effort levels as the wages in excess of fair levels. Excluding workers who receive exactly their fair wages yields an even more extreme asymmetry and the number of participants with wages above fair levels becomes smaller.

Table 8 – Estimates of the Effect of Wage Fairness on Effort from Tobit Regressions (II) and (III)

	(1)	(2)	(3)	(4)
Immediate Workers				
Gap Below Fair ($\beta_{imm-below}$)	0.257	0.077	0.303	0.116
	(0.321)	(0.288)	(0.352)	(0.309)
Gap Above Fair ($\beta_{imm-above}$)	-0.395	-0.411	-0.434	-0.412
	(0.307)	(0.309)	(0.335)	(0.336)
Wage (λ_{imm})	1.093***	1.142****	1.204***	1.273****
	(0.331)	(0.263)	(0.363)	(0.300)
Above Fair ($\delta_{imm-above}$)	-0.750	-0.492	-0.886	-0.685
	(0.554)	(0.458)	(0.596)	(0.499)
Delayed Workers				
Gap Below Fair ($\beta_{del-below}$)			-0.186	-0.071
			(0.685)	(0.649)
Gap Above Fair ($\beta_{del-above}$)			-0.530	-0.961
			(1.087)	(0.983)
Wage (λ_{del})			0.672	0.764
			(0.709)	(0.665)
Above Fair ($\delta_{del-above}$)			0.564	0.165
			(1.033)	(0.971)
Delay (γ)			0.873	1.407
			(2.731)	(2.491)
Constant	-1.055	1.665	-1.304	0.788
	(0.996)	(2.055)	(1.098)	(1.695)
Demographics	No	Yes	No	Yes
Pseudo-R^2				
N	60	60	116	116

Standard errors are in parentheses. P-values: *0.10, **0.05, ***0.01, ****0.001.

fair wage gap of -1 or +1 EUR (where the actual wage is 1 EUR less or more than the fair wage), by conducting separate F-tests of the restrictions that $\beta_{imm-above} + \lambda_{imm} = 0$ and $\beta_{imm-below} + \lambda_{del} = 0$. The F-test rejects that there is no reciprocity when the wage is below the fair wage by 1 EUR (p-value is always <0.001), and is close to rejection when the wage is above the fair level by 1 EUR (p-value is 0.080 without demographics, and 0.038 including demographics). Second, we test whether there is reciprocity when the fair wage gap is -2 or +2 EUR, with separate F-tests of the restrictions that $\beta_{imm-above} + \lambda_{imm} = 0$ and $\beta_{imm-below} + \lambda_{del} = 0$. The F-test rejects that there is no reciprocity when the wage is below the fair wage by 2 EUR (p-value is always <0.01), and does not reject it when the wage is above the fair wage by 2 EUR (p-value is 0.628 without demographics and 0.595 with demographics). Moreover, there is reciprocity for workers paid exactly their fair wage as the coefficient of wage is positive and significant (p-value is always <0.01).

How important is individual heterogeneity in perceptions of the fair wage for the asymmetry in short-term effort decisions in response to relatively low vs. high wages? We consider this by estimating equation (II) without the wage variable and restricting the fair wage for all immediate workers to be at the midpoint of the wage range, 2 USD. Reciprocity for wages below and above the midpoint is positive and significant with demographics, and close to significant without demographics (p-value below is 0.078 without demographics and 0.007 with demographics; p-value above is 0.052 without demographics and 0.014 with demographics). The coefficient of reciprocity below 2 USD is roughly six times as large. While this difference between coefficients is qualitatively similar to the results obtained when taking into account individuals' indicated fair wages, we are more likely to find significant reciprocity above the fair wage with this technique than when using self-reported fair wage. This indicates that taking the heterogeneity of views about fairness into account isolates the asymmetry in reciprocal behavior more clearly. The gain of our method in terms of explanatory power, however, is modest (pseudo-R^2 is 0.126 vs. 0.115 without demographics and 0.211 vs. 0.202 with demographics). We also note that a kink similar to the one in Tucson-H is observed by taking the middle point of the range as the fair wage in Tilburg-L, but not in Tilburg-H.

Our fourth hypothesis states that the asymmetry between negative and positive reciprocity would be the same for immediate and delayed workers. To examine this question, we return to the lower part of Table 7, which shows the correlations between the fair wage gap and effort, for both below and above the fair wages, and for immediate and delayed workers. Whereas for immediate workers the wage gap-effort correlation is stronger and only significant below fair wages (correlation is 0.460 below vs. -0.073 above fair levels), for delayed workers the correlations are more similar in size and insignificant both below and above fair wages (correlation is 0.211 below vs. -0.057 above fair levels; p-values from Spearman's rank correlation tests are 0.263 and 0.945).

To consider the asymmetry between negative and positive reciprocity in the short- and long-term in greater detail, we use the following Tobit regression (*III*), which is similar to regression number (II), but includes delayed workers and allows for different coefficients for the two types of workers:

$$\boldsymbol{e_j = \alpha + \gamma \times D_j + \delta_{imm-above} \times z_j + \beta_{imm-above} \times (wage_j - fairwage_j) \times z_j \times (1 - D_j) + \beta_{imm-below} \times (wage_j - fairwage_j) \times (1 - z_j) \times (1 - D_j) + \lambda_{imm} \times wage_j \times (1 - D_j) + \delta_{del-above} \times D_j \times z_j + \beta_{del-above} \times (wage_j - fairwage_j) \times z_j \times D_j + \beta_{del-below} \times (wage_j - fairwage_j) \times (1 - z_j) \times D_j + \lambda_{del} \times wage_j \times D_j + \varepsilon_j} \qquad (III)$$

The estimates are presented in columns 3 and 4 of Table 8. The coefficient $\beta_{del-above}$ has a large negative coefficient, suggesting that the wage-effort relationship is also less steep for wages above the fair level for delayed workers. This is only partly confirmed by separate F-tests, which come close to rejecting the restrictions that $\beta_{imm-above} + \lambda = 0$ with demographics (p-value is 0.196 without demographics and 0.075 including demographics), and does not reject the restrictions that $\beta_{imm-below} + \lambda = 0$ (p-value is 0.917 without demographics and 0.872 including demographics). This means that, if anything, there is a positive wage-effort relationship that is close to significant at wages below the fair wage, but no significant relationship at wages above it. We use an F-test to evaluate whether the asymmetry is the same for immediate and delayed workers by testing the restriction that $\beta_{imm-above} - \beta_{imm-below} = \beta_{del-above} - \beta_{del-below}$. The F-test is not significant (p-value is 0.781 without demographics and 0.783 with demographics).

Furthermore, we can test for differences in the average effort chosen in the short- and long-term for wages below and far below the fair wage, as presented in Table 7. The average effort for wages falling short of the fair wage by more than 0.50 EUR is 0.859 (0.179) for immediate workers and 1.456 (0.213) for delayed workers, which is significantly higher using a t-test (two-sided p-value is 0.036).[23] The average effort for wages below the fair wage by any amount increases from 1.192 (0.179) for immediate to 1.503 (0.206) for delayed workers, but the difference is not significant using a t-test (two-sided p-value is 0.259).[24] Only effort levels for wages far enough below the fair wage are greater in the long- than in the short-term.

[23] The increase is also significant with a one-sided Mann-Whitney test (two-sided p-value is 0.064).
[24] Mann-Whitney test p-value is 0.379.

Results 3 and 4 describe our findings concerning hypotheses 3 and 4.

Result 3*: Immediate workers reciprocate higher wages with more effort when wages are below or equal to self-reported fair wages. Reciprocity above fair wages is weaker, and it is only significant when wages are slightly above fair levels.*

Result 4*: For delayed workers, reciprocity is close to significant only when wages are slightly below fair levels. When wages are above or substantially below fair levels, there is no reciprocity. Moreover, compared to immediate workers, delayed workers choose significantly higher effort for wages that are far below fair levels.*

4.3 Exploratory Analysis

In addition to evaluating our prior hypotheses, we consider two issues. The first is whether there is a relationship between emotions and effort provision, and whether this differs between the two types of worker. Specifically, overall emotional valence and anger are our prime concerns, because they have been associated with reciprocity in other contexts.[25] The second issue is whether employers correctly anticipate the reciprocal wage-effort relationship.

The summary statistics regarding the change in emotions of workers when they learn their wage in Tucson-H are presented in Table 9, separated by types of worker. The table shows the average intensity of each emotion prior to learning the wage and the change in intensity when learning the wage. We also show the change in intensity corresponding to the 1st quartile and the 3rd quartile change. We leave out 51% of the participants because their videos do not conform to what the FaceReader software requires (the face in full view) or because of recording problems.[26] This leaves 57 observations. The table reveals some interesting patterns. The first is that among both immediate and delayed workers, at least 25% experience an increase, and at least 25% undergo a decrease, in the strength of each emotion. The second is that anger is the emotion with the largest average increase for both types of workers. The third is that

[25] Emotional valence is calculated by taking the intensity of happiness and subtracting the intensity of the strongest negative emotion, i.e. the maximum intensity score of anger, disgust, fear, and sadness.

[26] Many participants started the experiment seated such that the camera could record their whole face, but then leaned down or slouched until a part of their face was out of view. A few were covering their face with their hands at the crucial moment when they received their wage. A few cameras also stopped recording because of malfunctions.

the 3rd quartile change in the intensity of anger is larger for immediate than delayed workers, suggesting that the immediate workers experience a greater level of anger.

Table 9 – Changes in Emotions of Workers when Learning their Wage

	Immediate Workers				*Delayed Workers*			
	Prior	Change			Prior	Change		
	Average	Average	1st Quart.	3rd Quart.	Average	Average	1st Quart.	3rd Quart.
Valence	-0.245	-0.011	-0.162	0.156	-0.182	-0.054	-0.176	0.078
	(0.221)	(0.286)			(0.241)	(0.217)		
Neutrality	0.356	-0.003	-0.075	0.121	0.375	0.029	-0.070	0.173
	(0.209)	(0.242)			(0.169)	(0.172)		
Happiness	0.028	0.000	-0.008	0.013	0.035	-0.017	-0.024	0.005
	(0.041)	(0.046)			(0.062)	(0.063)		
Surprise	0.158	-0.050	-0.065	0.070	0.126	-0.012	-0.081	0.027
	(0.209)	(0.169)			(0.144)	(0.230)		
Sadness	0.096	0.000	-0.025	0.002	0.095	-0.011	-0.094	0.043
	(0.185)	(0.225)			(0.171)	(0.102)		
Anger	0.120	0.055	-0.056	0.150	0.090	0.042	-0.025	0.109
	(0.152)	(0.230)			(0.144)	(0.110)		
Fear	0.0148	-0.003	-0.012	0.000	0.005	0.017	-0.001	0.007
	(0.027)	(0.038)			(0.012)	(0.073)		
Disgust	0.123	-0.048	-0.088	0.016	0.085	0.010	-0.051	0.029
	(0.188)	(0.176)			(0.157)	(0.183)		
N	29				28			

Each emotion has a separate intensity score from 0 to 1. To calculate a change, we use the difference between the average in the 45 seconds after learning the wage and the average in the prior 45 seconds. Participants were included only if the software could capture their emotions at least 40% of the time both prior and after.

In the upper part of Table 10, we present the average effort chosen for different ranges of change in anger levels, as well as the average change in anger levels for different wage ranges. We present these separately for immediate and delayed workers.[27] Two main features are visible. First, immediate workers who become angrier choose lesser effort levels. Second, immediate workers receiving a lower wage tend to become angrier. These two features do not appear to be present for delayed workers.

In the lower part of Table 10, we show the correlations between anger and effort, and between wages and anger, as well as the one-sided p-value of the non-parametric Spearman's rank correlation test. For

[27] We present a more detailed table in Appendix B.

Table 10 – Average Effort for Different Ranges of Anger, and Average Anger for Different Ranges of the Wage

	Immediate Workers		*Delayed Workers*	
Anger	Average Effort	N	Average Effort	N
[-1; -0.05 [	1.763 (1.162)	9	3.333 (1.155)	3
[-0.05; +0.05]	1.77 (1.321)	10	1.631 (1.401)	16
] +0.05; +1]	1.109 (1.300)	11	1.8 (1.487)	9
Wage	Average Anger	N	Average Anger	N
0-1	0.262 (0.169)	4	0.038 (0.088)	4
1.1-2	0.115 (0.300)	5	0.104 (0.101)	3
2.1-3	-0.016 (0.262)	12	0.033 (0.094)	13
3.1-4	0.018 (0.044)	8	0.036 (0.152)	8

Correlation between Anger and Effort, and between Wage and Anger

	Anger & Effort	N	Wage & Anger	N
Immediate Workers	-0.315 [0.055]	29	-0.375 [0.040]	29
Delayed Workers	0.019 [0.267]	28	-0.072 [0.112]	28

Standard errors are in parentheses. We indicate the one-sided p-values from Spearman's rank correlation tests in square brackets.

immediate workers, the correlation between anger and effort is negative and almost significant (the correlation is -0.315, the one-sided p-value is 0.055), and the correlation between wage and anger is negative and significant (correlation is -0.375, one-sided p-value is 0.040). For delayed workers, the correlation between anger and effort is close to zero and not significant (correlation is 0.019, one-sided p-value is 0.267), and the correlation between wage and anger is negative and not significant (correlation is -0.072, one-sided p-value is 0.112). The tests thus confirm that immediate workers become significantly angrier the lower the wage they receive, and for them, greater anger is associated with lower subsequent effort (although only at slightly above significance levels). For delayed workers, low wages do not generate anger and anger is not associated with subsequent effort. Moreover, delayed workers who become angrier (with an increase in anger level above 0.05, the third category in Table 10) choose a significantly or near significantly higher effort than immediate workers (one-sided p-value from t-test is 0.0610; one-sided p-value from Mann-Whitney test is 0.041).

We can analyze the correlations between anger and effort, and between the fair wage gap (wage – fair wage) and anger, for the participants who receive a wage below their self-reported fair wage. For immediate workers, the negative correlation between anger and effort is stronger and significant (correlation is -0.472, Spearman one-sided p-value is 0.041, N=18), and the same is true for the negative correlation between the fair wage gap and anger (correlation is -0.569, Spearman one-sided p-value is

0.014, N=18). For delayed workers, the correlation between anger and effort is negative and insignificant (correlation is -0.239, Spearman one-sided p-value is 0.136, N=19), and the negative correlation between the fair wage gap and anger is also insignificant (correlation is -0.106, Spearman one-sided p-value is 0.666, N=19). Therefore, the relationships between the fair wage gap and anger and between anger and subsequent effort are qualitatively stronger for both types of workers paid below their fair wage than when we consider the absolute wage and all workers, although the significance levels are similar.[28] [29]

To test parametrically whether changes in emotions are related to effort, while accounting for the censoring of effort at 0, we estimate the following Tobit regression:

$$\boldsymbol{e_j = \alpha + \pi \times M_j + \varepsilon_j} \tag{IV}$$

where M_j includes either changes in the intensity of emotional valence or individual emotions. The regression is estimated for immediate and delayed workers' effort separately, and the estimates are given in Table 11. We include the demographic controls in the estimation. Columns 1 – 6 report estimates for immediate workers, while 7 – 9 contain those for delayed workers. Removing the demographic controls from the regression does not affect the significance of our estimates or the size of the anger coefficient.

Column 2 presents the estimate of anger when it is the only emotion in the regression. The coefficient of anger is negative and significant (p-value is 0.002). The coefficient is approximately -3.5. A change corresponding to the 3rd quartile of change in anger (0.15) therefore results in a 0.5 reduction in effort. This is large relative to the average effort of 1.4 (s.e. 1.2) for immediate workers. By comparing columns 2 and 3, we see that anger explains less variation than the wage itself (pseudo-R^2 is 0.164 when including anger and excluding wage and 0.215 when excluding anger and including wage). Column 6 presents the estimates of each individual emotion when they are all included in the regression. The coefficient of anger is still negative and significant (p-value is 0.040), whereas none of the other individual emotions or overall valence has a significant coefficient.

---

[28] For immediate workers, the correlations between anger and effort, and between fair wage gap and anger, for workers paid less than fair levels are not significantly different than the correlations between anger and effort and between wage and anger for all immediate workers in Table 10 (Fisher's r to z transformation, two-sided p-values are 0.562 and 0.435).

[29] There are relatively few immediate and delayed workers who receive exactly or more than the wage that they consider fair, making it difficult to draw conclusions about the correlations between anger and effort chosen, and between anger and the fair wage gap, in this wage range. Nevertheless, for these workers, two correlations are positive and two are near zero. None is significant.

Table 11 – Estimates from Tobit Regression (IV) of Effort on Emotions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Immediate Workers						*Delayed Workers*		
Valence	0.932						1.139		
	(1.378)						(2.477)		
Anger		-3.614***		-2.089*	-7.301	-2.964**		2.343	1.165
		(1.062)		(1.061)	(5.028)	(1.382)		(4.754)	(6.741)
Sadness						1.767			-4.296
						(1.731)			(4.893)
Disgust						0.847			0.468
						(2.017)			(3.763)
Fear						7.747			-9.589
						(5.952)			(7.529)
Surprise						1.472			-2.690
						(1.759)			(3.629)
Happiness						-1.130			1.221
						(6.323)			(9.997)
Wage			1.040****	0.832***	0.696*				
			(0.271)	(0.292)	(0.343)				
Int. Wage and Anger					2.112				
					(1.886)				
Age	-0.108	-0.039	-0.210	-0.164	-0.160	0.074	-0.103	-0.134	-0.185
	(0.192)	(0.149)	(0.179)	(0.160)	(0.172)	(0.166)	(0.133)	(0.126)	(0.178)
Man	1.279**	1.561**	0.835*	1.135**	1.055**	1.486**	0.842	0.918	0.985
	(0.570)	(0.581)	(0.433)	(0.485)	(0.504)	(0.612)	(1.104)	(1.106)	(1.443)
Friend	1.138	1.244	1.746***	1.572**	1.502**	1.975	-0.823	-1.018	-0.536
	(1.125)	(0.917)	(0.596)	(0.665)	(0.717)	(1.230)	(0.536)	(0.624)	(0.910)
American	0.237	0.598	0.240	0.441	0.516	0.571	-0.944	-1.107	-1.367
	(0.695)	(0.620)	(0.683)	(0.647)	(0.633)	(0.596)	(0.714)	(0.779)	(0.870)
Constant	2.741	1.133	2.145	1.618	1.954	-1.063	4.607	5.232*	6.551*
	(3.904)	(3.004)	(3.777)	(3.364)	(3.460)	(3.274)	(2.691)	(2.536)	(3.301)
Pseudo-R^2	0.080	0.164	0.215	0.251	0.261	0.197	0.030	0.031	0.081
N	29						28		

Each emotion has a separate intensity score from 0 to 1. Standard errors are in parentheses. P-values: *0.10, **0.05, ***0.01, ****0.001.

Including wage and anger in the same regression yields a significant wage coefficient, and renders anger borderline insignificant (column 4; p-value is 0.061). This suggests that low wages have a direct effect on effort, and some of the relationship between anger and effort in other specifications is a consequence of the effect of wages on the anger level. We include the term for the interaction of wage and anger in column 5. The wage is marginally insignificant (p-value of 0.055), and the anger and the interaction terms are not significant. The interaction is difficult to interpret since, as we discuss later, the wage influences anger. Nevertheless, participants who become angrier when learning that they receive a low wage increase effort more in reaction to a wage increase (F-tests do not reject the restriction 1 $\times$ [wage] - 0.05 $\times$ [interaction anger] = 0, which corresponds to the 1st quartile change in anger, with p-value of 0.155, and reject the restriction 1 $\times$ [wage] + 0.15 $\times$ [interaction anger] = 0, which corresponds to the 3rd quartile change in anger, with p-value of 0.004).

As an additional examination not shown in Table 11, we regress immediate effort on each individual emotion separately. Only sadness has a significant coefficient (coefficient is -3.3 (s.e. 1.2), p-value is 0.009). The coefficient of sadness is similar in magnitude to the one for anger, but is positive in sign. This means that an increase in sadness indicates an increase in effort. This might be due to the correlation between anger and sadness (correlation is -0.35). A regression including both anger and sadness yields a significant coefficient for anger only. A regression that includes anger, sadness, and the wage results in an insignificant anger coefficient and a significant sadness coefficient, but only when demographics are included in our specification.[30]

The above relationships do not appear for delayed workers. Neither emotional valence nor any of the individual emotions have a significant coefficient in columns 7-9.[31]

We also study the effect of wages on emotions. To do this, we regress each emotion on wage individually using OLS. The estimates are shown in Table 12. The wage has a significant negative effect on the anger level of immediate workers (p-value is 0.005) and no significant effect for delayed workers.[32] Furthermore, we find that the emotional valence of both types of workers is significantly increased by a higher wage. The findings are summarized as our fifth result.

Result 5*: For immediate workers, (1) anger is negatively associated with subsequent effort, and (2) lower wages lead to greater anger. For delayed workers, neither of these two relationships hold.*

[30] While a positive relationship between sadness and effort might be surprising at first glance, it is consistent with a common theme in psychology. A large body of work emphasizes the role of empathy in generating altruistic behavior (see Batson 2011 for a recent overview). Our analysis finds that the wage is not responsible for the sadness experienced. It could be, then, that workers who are able to feel how the employer would feel if effort were low, are the ones giving more. Sadness may reflect the expression of this empathy.
[31] Again, excluding demographics does not make a difference.
[32] Excluding demographics yields a similar coefficient for anger of immediate workers of -0.091 (s.e. 0.032; p-value < 0.001), and does not change the lack of significance of anger for delayed workers.

Table 12 – Effect of Wage on Emotional Changes in Workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Emotion	Anger	Sadness	Disgust	Fear	Surprise	Happiness	Valence
Immediate workers							
Wage	-0.111***	0.009	-0.048	0.004	0.037	0.010	0.121**
	(0.036)	(0.034)	(0.031)	(0.008)	(0.046)	(0.007)	(0.055)
Constant	-0.352	0.794	-0.248	0.042	0.385	-0.062	-0.294
	(0.450)	(0.511)	(0.542)	(0.101)	(0.462)	(0.139)	(0.538)
Adj. R^2	0.159	0.171	0.037	-0.065	-0.076	-0.049	0.269
N	29						
Delayed workers							
Wage	-0.007	-0.011	-0.067*	-0.010	-0.028	0.015*	0.094**
	(0.020)	(0.021)	(0.035)	(0.008)	(0.039)	(0.007)	(0.037)
Constant	-0.041	0.011	-0.197	0.076	0.527**	0.032	0.141
	(0.114)	(0.092)	(0.373)	(0.060)	(0.240)	(0.044)	(0.329)
Adj. R^2	-0.129	-0.030	0.095	-0.132	-0.012	0.058	0.161
N	28						

Each emotion has a separate intensity score from 0 to 1. Demographic controls are included. Standard errors are in parentheses. P-values: *0.10, **0.05, ***0.01, ****0.001.

The emotions of employers when they choose their wage are unrelated to their choice of wage. We find no significant correlation between individual emotions and the choice of wage, using Spearman tests.

To explore whether employers anticipate the wage-effort relationship for each type of worker correctly, we regress the effort prediction of employers on the wage that the employer offered. We use a Tobit regression. The estimates are shown in Table 13.

In Tucson-H, employers offering a higher wage predict higher effort by immediate workers, but not by delayed ones. In Tilburg-H, employers offering a higher wage predict a higher effort for both groups of worker. This result does not establish a causal link between wages and effort predicted, since it is consistent with the existence of reciprocal behavior on the part of the employer, but also with the anticipation on the part of the employer of reciprocal behavior by the employee. The significant coefficient on the variable European suggests that there may be cultural differences in employer expectations.

Table 13 – Regression of Employers' Effort Prediction on the Wage Offered

	Tilburg-H		Tucson-H	
Effort Predicted	*Immediate*	*Delayed*	*Immediate*	*Delayed*
Wage	0.883****	0.718****	0.432***	0.221
	(0.919)	(0.121)	(0.128)	(0.188)
Age	0.013	-0.013	0.042	0.076*
	(0.036)	(0.038)	(0.029)	(0.040)
Man	-0.344	-0.266	-0.177	-0.320
	(0.248)	(0.245)	(0.237)	(0.335)
Friend	0.257	-0.059	0.466*	0.561
	(0.208)	(0.248)	(0.257)	(0.574)
American			0.442	0.472
			(0.275)	(0.337)
European	0.561***	0.497**		
	(0.195)	(0.222)		
Constant	-0.446	0.454	-0.120	-0.386
	(0.861)	(0.914)	(0.610)	(0.910)
Pseudo-R^2	0.498	0.323	0.129	0.092
N	33	33	65	65

Standard errors are in parentheses. P-values: *0.10, **0.05, ***0.01, ****0.001.

Our sixth result, which summarizes our findings regarding employer predictions, is the following:

Result 6*: Employers' effort predictions are consistent with anticipating the existence of a positive wage-effort relationship for immediate workers. For delayed workers, this relationship is weaker and inconsistent.*

5. Conclusion

Akerlof and Yellen (1988, 1990) propose the Fair Wage-Effort hypothesis, a notion that workers shirk in retaliation for wages below the level that they perceive as fair. They propose that the process whereby this occurs is that low wages generate anger and in turn, anger leads to low effort. Furthermore, they assume that there is no reciprocity for wages above the fair level. The existence of this asymmetry has been supported by the field experiment reported by Cohn et al. (2014). They found that the reciprocal relationship between wages and effort held only for wages below what workers considered fair, and that the relationship was stable over time during a temporary employment relationship. Here, we consider whether the same relationship is observed in a simpler one-shot, but delayed, interaction, which allows us to clearly isolate any reciprocal behavior from other decisions. We also consider the relationship between wages, anger, and effort.

Some features of our experiment allow us to evaluate other patterns in the Akerlof and Yellen model. The model postulates an emotional mechanism as the source of the reciprocal wage-effort relationship. Emotions may be short-lived, and we consider whether the asymmetric relationship in their model persists over time. Comparing worker decisions on the same day and one month after learning their wage allows us to test whether workers' reciprocation of employer wage decisions would decrease with a long delay between the time workers learn their wage and the time they choose their effort. Eliciting self-reported fair wages aimed to test whether reciprocity is stronger for wages below vs. above fair levels, and whether any observed asymmetry is stable over time. We also consider whether a long delay affects employers' expectations and workers' emotional reactions.

We observe that, consistent with the models of Akerlof and Yellen, lower wages are reciprocated with lower effort, but more extensively so when they are less than or equal to a level the worker considers fair. This asymmetric reciprocal relationship is very pronounced for those making their decisions on the same day as they learn their wage, and weaker for those choosing their effort one month later. In the short term, workers decrease their effort the farther their wage is below the wage they deem fair, and increase their effort for a higher wage only until their wage is slightly above the fair wage. In the long term, workers exhibit some increase in effort in response to a higher wage, but only when their wage is slightly below the fair wage, as if they condition their effort on whether the fair wage threshold has been met or not. Specifically, workers offered wages far below their fair levels choose greater effort after one month than on the same day. While the link between wages and effort is sharper in the short-term, overall effort is similar over the two time horizons. Our results are similar to Cohn et al. (2014)'s finding that reciprocity is restricted to workers who feel paid less than a fair wage, although we also find that the asymmetry in reciprocity around the fair wage is less pronounced after a long time span.

We obtain some additional results. First, employers' predictions of workers' effort are consistent with a reciprocal wage-effort relationship in the short-term, but the relationship is less consistent in the long-term. Thus, employers seem to demonstrate an awareness that the reciprocal relationship between wage and effort exists in the short term, and tends to dissipate in the long term.

Second, anger on the part of workers is associated with low effort decisions undertaken within three hours. However, this relationship does not appear one month afterwards. Low wages provoke anger in immediate workers, but less anger in their delayed colleagues. These relationships are at least as strong when we use the difference between actual and fair wages instead of absolute wages. These patterns between wages, anger, and effort in the short term are consistent with the negative reciprocity assumed in the Fair Wage-Effort hypothesis.

It is well-known from previous studies that experienced anger can affect decisions taken immediately. Recall that we measure anger here at the time a worker learns of her offer, rather than at the time she chooses her effort. However, it is interesting that anger is stronger when one must react to a decision in the near rather than the more distant future. It is noteworthy that the change in anger when learning one's wage is a good indicator of subsequent effort even within three hours, although not as good an indicator as the wage itself. We also provide some evidence that wages might interact with anger. A lower wage results in less effort for those immediate workers who become angrier at the time of learning their wage.

Anger is the only emotion, of those we are able to measure, that is correlated to both wages received and effort chosen. Sadness is unrelated to wage, though positively correlated with effort. This may be evidence of sadness reflecting empathy, and feelings of empathy leading to higher effort with the purpose of benefiting the other player. Overall emotional valence, though negatively affected by low wages, is not associated with effort on the part of either immediate or delayed workers.

As discussed in the introduction, the previous literature has yielded mixed results with regard to whether the reciprocal relationship between wage and effort is durable. Our reading of the literature (Gneezy and List, 2006; Bellemare and Shearer, 2009; Hossain and List 2012; Kube et al., 2012, 2013; Cohn et al., 2014; Gilchrist et al., 2016) is that the persistence of the relationship depends on context, prior and future interaction between workers and employers, and culture. It not implausible that some cultures have a stronger norm of rewarding good deeds and punishing bad ones.[33] In our almost completely decontextualized laboratory setting, we observe that the wage-effort relationship weakens over time in

[33] For cross-cultural reviews, see Oosterbeek, Sloof, and van De Kuilen (2004) on Ultimatum Games; Cardenas and Carpenter (2008) on Trust and Ultimatum Games; and Henrich, McElreath, Barr, Ensminger, Barrett, Bolyanatz, Cardenas, Gurven, Gwako, Henrich, Lesorogol, Marlowe, Tracer, and Ziker (2006), Gächter and Herrmann (2009) and Henrich et al. (2010) for the willingness to punish.

our Dutch population more than in our American population. Nonetheless, average effort within each population is the same when effort decisions are taken on the same day or delayed by a month.

It may be that workers are more forgiving of low wages after a long cooling off period, and this might have to do with the dissipation of the anger associated with receiving a low wage. More generally, a reinforcement of negative reciprocity through a channel of anger would be consistent with other findings that negative reciprocity is stronger than positive reciprocity (e.g. Offerman, 2002; Charness, 2004; Engelmann and Ortmann, 2009; Kube et al., 2013; Cohn et al., 2014). Alternatively, it may be that workers learning now that they will receive low wages in the future discount their importance, adapt to them over time, and therefore do not punish the employer once they finally receive them. This can explain why we observe less anger at the low wages from employees who know they will receive their wage only one month later. This is consistent with reference-point models where one adapts to a lower reference wage (e.g. Kőszegi and Rabin, 2006).

Although it has been traditional to view emotions as outside the purview of economics, there is increasing recognition that they are important for decision making. Anger, for example, is the psychological mechanism proposed in Akerlof and Yellen (1990) to explain why workers would work less if they received less than what they believed is their fair wage, and has been modelled in psychological game theory (Battigalli et al. 2015). While including emotions in empirical work is a new and challenging avenue for economics, we believe that it might yield valuable insights into the mechanism whereby other-regarding preferences influence behaviour. Here we show that anger and not happiness is the emotion associated with effort, reaffirming the view that variations in effort reflect gradations in negative reciprocity. We believe that the direct measurement and tracking of emotions is useful in distinguishing between different models, such as the Partial Gift Exchange version of reciprocity and its Fair Wage-Effort extension.

One can ask how profitable it was for employers to offer a higher wage. In Tucson-H, offering a higher wage to both immediate and delayed workers was highly profitable. With a wage coefficient of 0.93 and 0.63 for the immediate and delayed workers, respectively, employers were increasing revenues by 2.33 USD (2.5*0.93) for a 1 USD increase in wage for immediate workers and by 1.58 USD (2.5*0.63) for delayed workers.[34] In Tilburg-H and Tilburg-L, with wage coefficients of 0.75, -0.37, 0.38, and 0.26, for short-term and long-term workers in each experiment, respectively, employers increased revenues by 1.88 EUR, -

[34] We use the wage coefficients in Table 5.

0.93 EUR, 0.95 EUR, and 0.65 EUR by offering a 1 EUR wage increase. This means that paying a higher wage to immediate workers resulted in greater or similar profits for employers, but that paying higher wages to delayed workers was profitable in only one of the three experiments. It is unclear whether these patterns in profitability extend to other parameters. Nonetheless, in all cases, once wages exceed the level a worker views as fair, employers lose profits by paying more. As such, the examination of the timing of the task at hand and employee views about fair wages can benefit employers trying to induce high effort from their employees.

References

Abeler, J., Altmann, S., Kube, S., & Wibral, M. (2010). Gift exchange and workers' fairness concerns: when equality is unfair. *Journal of the European Economic Association*, 8(6), 1299-1324.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics*, 97(4), 543-569.

Akerlof, G. A., & Yellen, J. L. (1988). Fairness and unemployment. *American Economic Review*, 78(2), 44-49.

Akerlof, G. A., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics*, 105(2), 255-283.

Batson, C. D. (2011). Altruism in humans. Oxford University Press, New York, NY, United States.

Battigalli, P., Dufwenberg, M., & Smith, A. (2015). Frustration and anger in games. Working Paper, Bocconi University.

Bellemare, C., & Shearer, B. (2009). Gift giving and worker productivity: Evidence from a firm-level experiment. *Games and Economic Behavior*, 67(1), 233-244.

Bolle, F., Tan, J. H., & Zizzo, D. J. (2014). Vendettas. *American Economic Journal: Microeconomics*, 6(2), 93-130.

Bosman, R., & Van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *Economic Journal*, 112(476), 147-169.

Breaban, A., & Noussair, C. N. (2014). Emotional state and market behavior. Working Paper, Tilburg University.

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics*, 2(3), 227-238.

Brandts, J., & Charness, G. (2004). Do labour market conditions affect gift exchange? Some experimental evidence. *Economic Journal*, 114(497), 684-708.

Brown, M., Falk, A., & Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, 72(3), 747-780.

Casoria, F., & Riedl, A. (2013). Experimental labor markets and policy considerations: incomplete contracts and macroeconomic aspects. *Journal of Economic Surveys*, 27(3), 398-420.

Cardenas, J. C., & Carpenter, J. (2008). Behavioural development economics: lessons from field labs in the developing world. *Journal of Development Studies*, 44(3), 311-338.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of labor Economics*, 22(3), 665-688.

Charness, G., Cobo-Reyes, R., Lacomba, J. A., Lagos, F., & Perez, J. M. (2016). Social comparisons in wage delegation: experimental evidence. *Experimental Economics*, 19(2), 433-459.

Charness, G., Fréchette, G. R., & Kagel, J. H. (2004). How robust is laboratory gift exchange?. *Experimental Economics*, 7(2), 189-205.

Charness, G., & Kuhn, P. (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of Labor Economics*, 25(4), 693-723.

Charness, G., & Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab?. In Ashenfelter, O., & Card, D., eds., Handbook of Labor Economics, 4(B). Elsevier, Amsterdam, The Netherlands.

Chentsova-Dutton, Y. E., & Tsai, J. L. (2010). Self-focused attention and emotional reactivity: the role of culture. *Journal of Personality and Social Psychology*, 98(3), 507-519.

Cohn, A., Fehr, E., & Goette, L. (2014). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, 61(8), 1777-1794.

Cohn, A., Fehr, E., Herrmann, B., & Schneider, F. (2014). Social comparison and effort provision: Evidence from a field experiment. *Journal of the European Economic Association*, 12(4), 877-898.

Cooper, D. J., & Kagel, J. H. 2016. Other regarding preferences: a selective survey of experimental results. In Kagel, J. H., & E. Roth, A. E., eds., Handbook of Experimental Economics, 2. Princeton University Press, Princeton, NJ, United States.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268-298.

Dreber, A., Fudenberg, D., Levine, D. K., & Rand, D. G. (2016). Self-control, social preferences and the effect of delayed payments. Working Paper, Washington University in Saint Louis.

Ekman, P., & Friesen, W. V. (1977). Facial action coding system. Consulting Psychologists Press, Palo Alto, CA, United States.

Ekman, P., & Friesen, W. V. (2003). Unmasking the face: a guide to recognizing emotions from facial clues. ISHK, Los Altos, CA, United States.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Rainer, K., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712-717.

Engelmann, D., & Ortmann, A. (2009). The robustness of laboratory gift exchange: a reconsideration. Working Paper, Charles University.

Fehr, E., & Falk, A. (1999). Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy*, 107(1), 106-134.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: experimental evidence. *Econometrica,* 65(4), 833-860.

Fehr, E., Goette, L., & Zehnder, C. (2009). A behavioral account of the labor market: the role of fairness concerns. *Annual Review of Economics*, 1, 355-384.

Fehr, E., Kirchler, E., Weichbold, A., & Gächter, S. (1998). When social norms overpower competition: Gift exchange in experimental labor markets. *Journal of Labor Economics*, 16(2), 324-351.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, 108(2), 437-459.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1996). Involuntary unemployment and non-compensating wage differentials in an experimental labour market. *Economic Journal*, 106(434), 106-121.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42(1), 1-34.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.

Gächter, S., & Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *Scandinavian Journal of Economics*, 104(1), 1-26.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518), 791-806.

Gächter, S., & Thöni, C. (2010). Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis. *Journal of Economic Behavior & Organization*, 76(3), 531-543.

Gadea, M., Aliño, M., Espert, R., & Salvador, A. (2015). Deceit and facial expression in children: the enabling role of the "poker face" child and the dependent personality of the detector. *Frontiers in Psychology*, 6(1089).

Gilchrist, D. S., Luca, M., & Malhotra, D. (2016). When 3+ 1> 4: Gift structure and reciprocity in the field. *Management Science*, 62(9), 2639-2650.

Gneezy, U., & List, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5), 1365-1384.

Grimm, V., & Mengel, F. (2011). Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Economics Letters*, 111(2), 113-115.

Guth, W., Konigstein, M., Kovács, J., & Zala-Mezõ, E. (2001). Fairness within firms: The case of one principal and multiple agents. *Schmalenbach Business Review: ZFBF*, 53(2), 82.

Habětínová, L., & Noussair, C. N. (2015). Charitable giving, emotions, and the default effect. Working Paper, Tilburg University.

Hannan, R. L., Kagel, J. H., & Moser, D. V. (2002). Partial gift exchange in an experimental labor market: impact of subject population differences, productivity differences, and effort requests on behavior. *Journal of Labor Economics*, 20(4), 923-951.

Hennig-Schmidt, H., Sadrieh, A., & Rockenbach, B. (2010). In search of workers' real effort reciprocity—a field and a laboratory experiment. *Journal of the European Economic Association*, 8(4), 817-837.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2), 73-78.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.-C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767-1770.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and Brain Sciences*, 33(2-3), 61-83.

Offerman, T. (2002). Hurting hurts more than helping helps. *European Economic Review*, 46(8), 1423-1437.

Hopfensitz, A., & Reuben, E. (2009). The importance of emotions for the effectiveness of social Punishment. *Economic Journal*, 119(540), 1534-1559.

Hossain, T., & List, J. A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, 58(12), 2151-2167.

Kessler, J. B. (2013). When will there be gift exchange? Addressing the lab-field debate with laboratory gift exchange experiments. Working Paper, University of Pennsylvania.

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

Kovarik, J. (2009). Giving it now or later: Altruism and discounting. *Economics Letters*, 102(3), 152-154.

Kube, S., Maréchal, M. A., & Puppe, C. (2012). The currency of reciprocity: gift exchange in the workplace. *American Economic Review*, 102(4), 1644-1662.

Kube, S., Maréchal, M. A., & Puppe, C. (2013). Do wage cuts damage work morale? Evidence from a natural field experiment. *Journal of the European Economic Association*, *11*(4), 853-870.

Lewinski, P., Fransen, M. L., & Tan, E. S. (2014). Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics*, 7(1), 1-14.

Lensberg, T., & van der Heijden, E. (1998). A cross-cultural study of reciprocity, trust and altruism in a gift-exchange experiment. Working Paper, Tilburg University.

List, J. A. (2006). The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions. *Journal of Political Economy*, 114(1), 1-37.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. *Health Psychology*, 24(4S), S49.

Maximiano, S., Sloof, R., & Sonnemans, J. (2007). Gift exchange in a multi-worker firm. *Economic Journal*, 117(522), 1025-1050.

Neo, W. S., Yu, M., Weber, R. A., & Gonzalez, C. (2013). The effects of time delay in reciprocity games. *Journal of Economic Psychology*, 34, 20-35.

Nguyen, Y., & Noussair, C. N. (2014). Risk aversion and emotions. *Pacific Economic Review*, 19(3), 296-312.

Oosterbeek, H., Sloof, R., & Van De Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7(2), 171-188.

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

Reuben, E., & Van Winden, F. (2008). Social ties and coordination on negative reciprocity: The role of affect. *Journal of Public Economics*, 92(1), 34-53.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755-1758.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, 49(2), 144-159.

Van Leeuwen, B., Noussair, C. N., Offerman, T. J. S., Suetens, S., Van Veelen, M., & Van de Ven, J. (2015). Predictably angry: Facial cues provide a credible signal of destructive behavior. *Management Science*, forthcoming.

Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20), 7398-7401.

Appendix A – Additional Table for Wage-Effort Relationships Below and Above Fair Wages

In Table 14, we describe the wage-effort relationships below and above the self-reported fair wage and distinguish between different wage and fair wage ranges.

Table 14 - Average Effort for Different Ranges of Fair and Actual Wage, for Immediate and Delayed Workers

Average Effort

Immediate Workers

Fair Wage \ **Wage**	0-0.5	0.6-1	1.1-1.5	1.6-2	2.1-2.5	2.6-3	3.1-3.5	3.6-4
0-0.5	.	.	.	.	.	.	.	1 [1]
0.6-1	.	.	.	.	.	.	.	0 [1]
1.1-1.5	.	.	.	.	.	.	.	.
1.6-2	.	0 [1]	1.6 [1]	0 [1]	0.2 [1]	1.1 [1]	2 [1]	.
2.1-2.5	.	0 [1]	0 [1]	.	2.05 [2]	.	3.5 [1]	2.1 [2]
2.6-3	.	1 [1]	0.1 [1]	0.45 [2]	1.3 [4]	1.97 [3]	0 [1]	2.33 [3]
3.1-3.5	.	0 [1]	.	1.55 [2]	0.9 [2]	1.34 [5]	.	3.03 [3]
3.6-4	0 [1]	0 [1]	.	1.16 [4]	0.55 [2]	2.37 [3]	3.17 [3]	3 [3]

Delayed Workers

Fair Wage \ **Wage**	0-0.5	0.6-1	1.1-1.5	1.6-2	2.1-2.5	2.6-3	3.1-3.5	3.6-4
0-0.5	.	.	.	.	.	.	.	.
0.6-1	.	.	.	.	.	.	.	.
1.1-1.5	.	.	.	.	.	.	.	.
1.6-2	.	.	.	0 [1]	0 [1]	.	2.2 [1]	.
2.1-2.5	.	.	.	.	0 [1]	4 [1]	2.1 [1]	1 [2]
2.6-3	.	1 [1]	.	2.5 [1]	2 [1]	.	.	3 [1]
3.1-3.5	0 [1]	.	2.5 [2]	0.75 [2]	2.05 [4]	2.32 [5]	2.5 [2]	4 [2]
3.6-4	.	1.2 [5]	1.8 [1]	0.63 [4]	1.25 [2]	1.83 [6]	1.5 [2]	1.75 [6]

The number of observations is indicated in square brackets.

We make two observations. First, average effort levels in response to low wages are higher for delayed than immediate workers, regardless of the reported fair wages. Second, in addition to greater effort levels resulting from higher wages, we observe that falling short of fair wages might decrease effort levels. Indeed, for wages below fair levels, effort levels are greater as wages increase, and, for wages above fair levels, effort levels are generally flatter.

Appendix B – Additional Table for Relationships between Wage and Anger and between Anger and Effort

In Table 15, we present the average effort chosen for different anger and wage ranges, separately for immediate and delayed workers. We can observe first that immediate workers who become angrier when learning their wage tend to choose less effort. Second, low wages generate more anger from immediate workers. These two relationships do not appear with clarity for delayed workers.

Table 15 – Average Effort for Different Anger and Wage Ranges, for Immediate and Delayed Workers

Average Effort

Immediate Workers

	Wage							
Anger	0-0.5	0.6-1	1.1-1.5	1.6-2	2.1-2.5	2.6-3	3.1-3.5	3.6-4
[-1; -0.3[	.	.	.	.	.	2.8 [1]	.	.
[-0.3; -0.2[	.	.	.	.	.	0.5 [1]	.	.
[-0.2; -0.1[	.	.	.	.	2.4 [1]	0.93 [3]	.	.
[-0.1; -0.05[	.	.	.	2 [1]	.	3.6 [1]	.	.
[-0.05; 0]	.	.	1.6 [1]	0.95 [2]	0.1 [1]	.	.	3 [3]
]0; 0.05]	.	.	.	.	.	.	1 [2]	3.1 [1]
]0.05; 0.1]	.	.	.	.	.	.	.	2 [2]
]0.1; 0.2]	.	0.5 [2]	.	.	1.85 [2]	.	.	.
]0.2; 0.3]	.	0 [1]	.	.	.	.	.	.
]0.3; 0.4]	.	.	.	.	.	.	.	.
]0.4; 1]	.	0 [1]	.	0 [1]	1.3 [1]	2.2 [1]	.	.

Delayed Workers

	Wage							
Anger	0-0.5	0.6-1	1.1-1.5	1.6-2	2.1-2.5	2.6-3	3.1-3.5	3.6-4
[-1; -0.3[	.	.	.	.	.	.	.	.
[-0.3; -0.2[	.	.	.	.	.	.	.	.
[-0.2; -0.1[	.	.	.	.	.	.	.	4 [1]
[-0.1; -0.05[	.	.	.	.	4 [1]	2 [1]	.	.
[-0.05; 0]	.	1.5 [2]	.	.	0 [1]	2.9 [4]	.	0.83 [3]
]0; 0.05]	0 [1]	.	4 [1]	.	0 [1]	2 [1]	.	1.5 [2]
]0.05; 0.1]	.	.	.	.	2 [1]	.	.	.
]0.1; 0.2]	.	2 [1]	.	0 [1]	0 [1]	1 [1]	.	.
]0.2; 0.3]	.	.	.	1 [1]	2.2 [1]	.	.	4 [2]
]0.3; 0.4]	.	.	.	.	.	.	.	.
]0.4; 1]	.	.	.	.	.	.	.	.

The number of observations is indicated in square brackets.

Appendix C – Instructions (for Tucson-H)

We present the instructions for the Tucson-H experiment. The instructions are very similar for the other two experiments.

Introduction

Welcome. The camera on top of your computer screen is currently turned on and you will be filmed for the entire session. The video will be seen only by the experimenters and will be destroyed once data have been gathered from it. The content of the video will not affect your payment. If you disagree with being filmed, you are free to leave at any time, and you will keep the show-up fee, but you will not receive additional payment.

In this experiment, the instructions are simple and if you read them carefully you can earn a considerable amount of money. Everything you earn is on top of a 5 dollar show- up fee and a 3 dollar participation fee. You are forbidden to communicate with other participants at any time. At the end the experiment, the earnings you have made will be paid to you by sending you cash or a check in the mail. Take a moment now to write down your name and your mailing address on the envelope in front of you. Please be careful, as we will use this envelope to send your payment.

This experiment takes around 40 minutes in the lab. In addition, depending on the role you are randomly assigned, you might have to take 5 minutes outside the lab to send an email, either in the 3 hours after the session today or in 4 weeks. If anything is unclear during the experiment, please raise your hand and the experimenter will come to you.

Roles

In this experiment, participants are randomly assigned, with equal chance, to 1 of the 3 roles: 1) employer, 2) employee choosing now or 3) employee choosing later. You will be placed in a group of 3 made of 1 employer, 1 employee choosing now, and 1 employee choosing later. The identity of the members of each group will be kept secret from all participants. Your role will be now revealed to you on your screen. Do not reveal your role to others and do not mention your role aloud to the experimenter.

General Instructions

This is an experiment about employer-employees interaction. First, the employer chooses a wage between 0 and 4 dollars for his/her 2 employees and asks them to choose a specific effort level between 0 and 4. The wage is the same for each employee. The employer cannot force the employees to choose a specific effort level. Second, the employees learn their wage and each employee chooses an effort level between 0 and 4. The employee choosing now will choose an effort level within 3 hours after the end of today's session and the employee choosing later will choose an effort level in 4 weeks. A higher effort level costs an employee more but benefits the employer more. Each unit of effort chosen costs the employee 25 cents and increases the earnings of the employer by 2.50 dollars. Employees receive the wage, no matter what effort level they choose. In addition, the employer is asked to make a prediction about the effort of each employee. Also, before learning his/her wage, each employee is asked to state the wage he/she thinks would be fair to receive.

Specific Instructions for the Employer

The employer chooses a wage between 0 and 4 dollars for each of the two employees. The wage can be chosen in any increment of 10 cents. The wage is the same for both employees. The employer has to pay the wage twice because there are two employees, one choosing an effort today and one choosing an effort in 4 weeks. The wages are costly for the employer and benefit the employee. The wages are deducted from the employer's earnings. The wage for one employee is added to that employee's earnings. The employer also asks for a specific effort level from each employee. The specific effort asked is the same for each employee. The employer cannot force an employee to choose a specific effort. Each employee is free to choose an effort level that is the same or different from the level that is asked. Each unit of effort chosen by an employee will cost him/her 25 cents and pay the employer 2.50 dollars. Your payment calculation is described in the section Payment Calculation for All Roles.

In addition, the employer is asked to predict the effort level of each employee. This pre- diction will not be revealed to any participant. Once both employees have chosen their effort level, the prediction for one employee will be randomly chosen. If this prediction is within 0.4 of the effort chosen by the employee, the employer will receive an additional 1 dollar.

The payment to the employer will be sent by mail on the day after the receipt of the effort level chosen by the employee choosing in 4 weeks. The employer has no email to send. If an employee does not choose an effort level, the employer is paid back the wage that was given to the employee.

Specific Instructions for the Employee Choosing Now

The employer will offer each employee a wage between 0 and 4 dollars and ask for a specific effort level between 0 and 4. The wage chosen is deducted from the employer's earnings. The employer cannot force an employee to choose a specific effort, he/she can only ask for it. The employee will then have to choose an effort between 0 and 4. The effort can be chosen in any increment of 0.1. The effort is costly for the employee and benefits the employer. Each unit of effort chosen by an employee will cost him/her 25 cents and pay the employer 2.50 dollars. Your payment calculation is described in the section Payment Calculation for All Roles.

The employee choosing now chooses his/her effort level within the 3 hours after the end of today's session. He/she does not choose the effort level during the session. He/she chooses the effort level by sending an email to the experimenter. The email address of the experimenter is written on the last page of these instructions. You can leave with the last page of the instructions after today's session. Thus you can also write anything you might need on the last page. The envelope with your payment will be mailed on the day after the receipt of your email. Please respect the 3 hour deadline as it will be enforced.

If you send your effort choice more than 3 hours after the end of this session, you will only be paid your show-up and participation fees.

In addition, before learning his/her wage, each employee will be asked what is the wage that he/she would consider fair to receive. The wage can be anything between 0 and 4 dollars, in any increment of 10 cents. This fair wage will not be revealed to any participant in the experiment. It cannot influence the wage that an employee receives. Please simply try to be truthful even if it does not influence your earnings.

Specific Instructions for the Employee Choosing Later

The employer will offer each employee a wage between 0 and 4 dollars and ask for a specific effort level between 0 and 4. The wage chosen is deducted from the employer's earnings. The employer cannot force an employee to choose a specific effort, he/she can only ask for it. The employee will then have to choose an effort between 0 and 4. The effort can be chosen in any increment of 0.1. The effort is costly for the employee and benefits the employer. Each unit of effort chosen by an employee will cost him/her 25 cents and pay the employer 2.50 dollars. Your payment calculation is described in the section Payment Calculation for All Roles.

The employee choosing later chooses his/her effort level only in 4 weeks after today's session. He/she does not choose the effort level during the session. He/she chooses the effort level by sending an email to the experimenter. The email address of the experimenter is written on the last page of these instructions. You can leave with the last page of the instructions after today's session. Thus you can also write anything you might need on the last page. The date in 4 weeks is also written on the last page. You will have 5 days to send your effort then, with the starting day being the date in 4 weeks. The envelope with your payment will be mailed on the day after the receipt of your email. Please respect the 5 day period beginning in 4 weeks. If you send your choice of effort before or after the 5 day period, you will only be paid your show-up and participation fees. An email reminder will be sent to you on the starting day in 4 weeks. The wage, the effort asked by the employer, and a copy of these instructions will be in the email reminder.

In addition, before learning his/her wage, each employee will be asked what is the wage that he/she would consider fair to receive. The wage can be anything between 0 and 4 dollars, in any increment of 10 cents. This fair wage will not be revealed to any participant in the experiment. It cannot influence the wage that an employee receives. Please simply try to be truthful even if it does not influence your earnings.

Payment Calculation for All Roles

Earnings in dollars for participants in each of the 3 roles are described by these formulas:

Employer: *11 − 2 ∗ wage + 2.5 ∗ (effort of employee choosing now + effort of employee choosing later)*

Employee choosing now*: 4 + wage − 0.25∗ effort of employee choosing now*

Employee choosing later: 4 + wage − 0.25∗ effort of employee choosing later

Note that the 5 dollar show-up fee, the 3 dollar participation fee, and the potential payment for the employer's predictions are not included. Please take a minute to study these formulas carefully as your earnings depend on understanding them.

Practice

During the experiment, you can choose only once. Before you start, please practice with 2 examples on the last page. The examples will not affect your earnings. When you have completed them, please raise your hand and the experimenter will verify your answers. This is the best time to ask questions to the experimenter to be sure that you under- stand. Please do not hesitate to ask. Your earnings depend on your understanding of the instructions.

At the end of the session, when you are done with the questionnaire, please stay seated and raise your hand. The experimenter will come to you and ask you to sign a sheet before you can leave. Please do not communicate with others during that time and leave your envelope at your seat. You are free to leave with the last page of the instructions, on which the email address is written.

You can leave with this page after today's session. Feel free to write anything on it.

Example 1: The employer chooses a wage of U, one employee chooses effort T, and the other employee chooses effort L. Do not replace the variables with numbers. Please write down the earnings for each role:

Employer:

Employee choosing effort T:

Employee choosing effort L:

Example 2: The employer chooses a wage of Z, one employee chooses effort Q, and the other employee chooses effort S. Do not replace the variables with numbers. Please write down the earnings for each role:

Employer:

Employee choosing effort Q:

Employee choosing effort S:

Date today: xx\xx\2016

Date in 4 weeks: yy\yy\2016

Wage can be anything between 0 and 4, in any increment of 0.1.

Effort level can be anything between 0 and 4, in any increment of 0.1.

Email to send choice of effort: nickolasgagnon@email.arizona.edu

Appendix D – Email Reminders for Delayed Workers

The email reminders for delayed workers are similar across the three experiments. There are three differences. First, participants are reminded of the non-binding effort request of the employer in Tucson-H and Tilburg-H, but not in Tilburg-L, as employers do not request an effort level in Tilburg-L. Second, participants in Tilburg-L were initially instructed to answer within a time interval of 24 hours, but this time interval was subsequently extended. Third, the third paragraph of the reminder for Tucson-H is not included in the other two experiments. This paragraph shows the range of the wage. The wage range, however, is always available in the attached instructions for each experiment.

Reminder for Tucson-H (original instructions attached)

Dear participant,

4 weeks ago, you were assigned to the role of **Employee Choosing Later** in an experiment at the Economic Science Lab of the University of Arizona. You had to send an email to this address with your choice of effort level 4 weeks after the session.

This is the **reminder** that the date 4 weeks after is today, **[date]**. You have now 5 days to **send your choice of effort**. These 5 days include today.

During the session, you were placed in a group of 3 participants consisting of an employer, an employee choosing now, and an employee choosing later. The employer had to choose a wage between 0 and 4 dollars, in any increment of 10 cents, for each employee. The wage was the same for both employees. The wages are deducted from the employer's earnings.

The employer offered you a wage of $[wage] during the session and asked for an effort level of [effort]. You now have to choose your effort level. Note that the employer cannot force you to choose a specific effort level.

You can choose an effort level between 0 and 4, in any increment of 0.1. Each unit of effort that you choose will cost you 25 cents, and pay the employer 2.50 dollars.

Payment Calculation (in dollars):

Employer: 11−2*Wage +2.5*(Effort of Employee Choosing Now + Effort of Employee Choosing Later)

Employee Choosing Now: 4 + Wage − 0.25*Effort of Employee Choosing Now

Employee Choosing Later: 4 + Wage − 0.25*Effort of Employee Choosing Later

If anything is unclear, please contact me. I also attach a **copy of the instructions** for the experiment. The envelope with your payment will be sent by mail on the day after the receipt of your email.

You must respect the deadline to be paid.

Reminder for Tilburg-H (original instructions attached)

Dear participant,

One month ago, you were attributed the role of **Employee Choosing Later** in an experiment. This means you had to send me an email with your choice of effort 4 weeks after the experiment.

This is a **reminder** that the date 4 weeks after is today, **[date]**. You have now 5 days to **send me your choice of effort**.

An Employer offered you a wage of [wage] euro during the experiment and asked you to make an effort of [effort]. You now have to choose your effort (note that the employer cannot force you to choose an effort).

Each effort unit that you choose will cost you 1/4 euro, and will increase the earnings of the Employer by 2.5 euro. You can choose an effort between 0 and 2.4, with a precision of 0.1 units.

Example of Payment Calculation:

To calculate the payment, consider that W is the wage chosen by the employer, A is the effort chosen by one of the employee, and B is the effort chosen by the other employee. In addition, there is a fixed fee of 4 euro for the employer and of 5 euro for each employee. The earnings are then described by these formulas (note that the fixed fees are included):

Employer: 4 − 2*W + 2.5*(A+B) + 4

Employee choosing effort A: W − A/4 + 5

Employee choosing effort B: W − B/4 + 5

If anything is unclear, please do not hesitate to email me.

I also attach a **copy of the instructions** for the whole experiment to this email,

Please respect the deadline to be paid.

Reminder for Tilburg-L (original instructions attached)

Dear participant,

one month ago, you were attributed the role of **Employee Choosing Effort Later** in an experiment. This means you had to send me an email with your choice of effort 30 days after the experiment.

This is a **reminder** that the date 30 days after is tomorrow, **[date]**. During the 24h of that day, you can **send me your choice of effort**. If you send your choice **before or later the 24h of that day, you will not be paid**.

An employer offered you a wage of [wage] euro during the experiment and you are asked to choose an effort level.

Each effort unit that you choose will cost you 1 euro, and will increase the payoff of the Employer by 2.5 euro. You can choose an effort level with a precision of 0.1 units. You cannot choose an effort level higher than the wage you are paid.

Example of payment calculation:

The employer has a wealth of 8 euro. He\She decides a wage that is paid to each of the two employees. Each participant receives a participation fee of 4 euro.

The wage chosen is W, one employee chooses effort level of E1 and the other employee chooses effort level of E2. Note that E1 and E2 are variables and not numbers for this example.

Payment of Employer: 8 - 2*W + 2.5*(E1) +2.5*(E2) + 4 euro

Payment of 1st Employee: W - (E1) + 4 euro

Payment of 2nd Employee: W - (E2) + 4 euro

If anything is unclear, please do not hesitate to email me.

I also attach a **copy of the instructions** for the whole experiment to this email,

Please respect the deadline to be paid.