We report the results from three experiments embedded in the same overarching design, which extends the Gift Exchange paradigm for the study of worker–employer relationships. We focus on the effect of the length of the delay, between the time at which workers learn their wage and when they choose an effort level, on the relationship between wage and effort. We compare effort choices made within a few hours with those made several weeks afterward. We find that the strength of the wage-effort relationship decreases over time, and this change appears to be driven by workers who receive low wages. (JEL C91, J33, M52)

I. 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-clearing level. A robust relationship of workers reciprocating higher wages with greater effort is observed. These findings have spawned a large literature (Brandts and Charness 2004; Brown, Falk, and Fehr 2004; Charness 2004; Fehr et al. 1998a; Fehr and Falk 1999; Fehr, Gächter, and Kirchsteiger 1997; Fehr, Kirchsteiger, and Riedl 1996, 1998b; Gächter and Falk 2002; Hannan, Kagel, and Moser 2002). 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 immediate behavior, in the sense that all wage and effort decisions are made within one laboratory session, which is typically under 2 hours in duration. Several recent studies report field experiments focused on whether the reciprocal relationship between wage and effort is durable over longer time horizons. Reciprocity, in the form of increased effort by workers, is commonly found within a few hours following a wage increase. In contrast, the few studies analyzing behavior over a longer timeframe provide some evidence that the tendency toward reciprocation can survive for at least 1 day. The results are consistent with the notion that reciprocity fades over horizons of days or weeks.

Specifically, several experimental studies have considered worker reciprocity over a time frame of several hours. Kube, Maréchal, and Puppe (2012) observe higher effort for the 3 hours they monitor after a one-time nonmonetary
gift, but not for a comparable monetary gift. Gilchrist, Luca, and Malhotra (2016) detect a similar effect for 4 hours following an unexpected wage increase, but not when the wage increase is expected. Kube, Maréchal, and Puppe (2013) find no reciprocal response to a wage increase, but do observe a sharp decrease in productivity following a wage decrease, over the 6 hours of their experiment. In a study spanning two work days, Gneezy and List (2006) observe that reciprocity lasts only for a few hours after an increase in workers’ hourly wage. A few studies have observed an effect for a longer time span, but these studies are not conclusive. Bellemare and Shearer (2009) find reciprocity lasting for 1 day after a one-off bonus payment, and thereafter workers’ effort comes back down to previous levels. Cohn, Fehr, and Goette (2015) report higher productivity following increases in hourly wages for the entire time horizon of their study, which was several days. However, they vary workers’ wages more than once, which may renew the stimulus. Hossain and List (2012) find an increase in productivity lasting for 4 weeks after a one-time monetary payment, though they rely on a relatively small sample of workers.

Those studies differ from each other in many aspects. The time horizons analyzed differ, ranging from a few hours to 4 weeks. Effort decisions are also 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. Furthermore, worker behavior might depend on the context (Gilchrist, Luca, and Malhotra 2016; Hennig-Schmidt, Sadrieh, and Rockenbach 2010; Kube, Maréchal, and Puppe 2012, 2013; List 2006), the perception of the fairness of the baseline wage (Cohn, Fehr, and Goette 2015; Fehr, Goette, and Zehnder 2009), prior history (Bellemare and Shearer 2009), or anticipated future interaction between worker and employer (Hossain and List 2012).

In the experiments reported here, we consider whether reciprocal behavior is long-term in nature, and we do so in a very simple one-shot design. This feature rules out any effect of past or anticipated future interaction. We compare the decisions of “short-term” workers choosing their effort within 3 hours of learning their wage, with those of “long-term” workers making their choice 4 weeks after receiving their wage information. Before long-term workers are required to make their effort decision, they are reminded of the wage decision their employer made 4 weeks previously. This is similar to the typical situation of an employee or contractor who learns her wage when hired, and is then reminded about her wage at the time she actually starts the job later on.

Our short-term delay represents the case of workers who learn their wage shortly before their work begins, and then choose how much effort to exert. Under the “hot vs. cold” terminology, both of our effort decisions are “cold” decisions, made after a period of reflection. The longer delay allows us to measure the dissipation of the reciprocity motive between the two time horizons. There is exactly one decision made by each worker. This allows us to clearly identify 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–worker relationships to long time horizons. The experiments are embedded in the same overarching design that allows scope for a reciprocal worker–employer relationship to emerge. The second experiment improves possible censoring issues faced by the first, and the third experiment employs a richer design that allows us to consider reciprocity from additional perspectives. To take advantage of the entire database (266 workers), we analyze our three experiments together whenever possible. In

---

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.

2. To conduct this experiment, it is imperative to remind long-term workers of the wage that they were offered previously. Otherwise, many would forget the wage that they were offered, and we could not be sure that their effort decisions were taken with the knowledge of the wage that they had received.

3. “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) for a discussion of differences in the two settings.

4. One can ask whether reciprocal behavior decreases beyond a certain threshold precisely because workers feel that they have reciprocated a past wage increase. While this is a very interesting question, we believe it is important to first ask whether reciprocity decreases over time when only a single decision is taken. This is because finding that reciprocity decreases in a repeated setting could simply reflect the fact that workers are less inclined to reciprocate with the passing of time.
all three experiments, we consider whether workers choose effort levels to reciprocate employers’ wage decisions when wages were set either less than 3 hours or 1 month previously. Subjects are divided into groups of three: one employer and two workers (a short-term and a long-term worker).\(^5\) The employer chooses one wage level that applies to both workers. The short-term worker then chooses an effort level after leaving the laboratory, but at most 3 hours afterward. The long-term worker submits the effort decision 4 weeks after the experimental session. At the time their decision is due, long-term workers receive a reminder informing them of the wage they received and the payoff structure in effect.\(^6\)

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 is designed to allow a test of a key part of a later refinement of the hypothesis, the Fair Wage-Effort hypothesis of Akerlof and Yellen (1988, 1990), in each of the two timeframes. 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. In the second model, Akerlof and Yellen (1990) assume that workers form a belief regarding what the fair wage is, and that they exhibit an asymmetric response to the wages about this fair wage. Wages below those that workers perceive as fair are punished with low effort, but those in excess of fair levels do not result in additional effort. Cohn, Fehr, and Goette (2015) recently obtained support for this second model, finding that the sustained reciprocity they uncover in their field experiment is driven by the subset of workers who felt that they were paid unfairly before a wage increase. That is, only those workers who felt that they were paid an unfair wage before increased their productivity in response to the increase, at least over the timespan of several days that they study. We test whether we observe reciprocity when wages are below fair wages (which we refer to as negative reciprocity) as well as when wages are above fair wages (which we call positive reciprocity), and whether one type of reciprocity dissipates more than the other with time.\(^7\)

One reason that reciprocity may be weaker in the long term than in the short term is the role of emotions. In justifying their Fair Wage-Effort hypothesis, Akerlof and Yellen (1990) propose that wages below those considered as fair trigger a reaction of anger on the part of workers, and in turn the feeling of anger triggers low effort. There exists empirical evidence in economics and social psychology that anger is correlated with immediate (over seconds or a few minutes) negatively reciprocal behavior (Ben-Shakhar et al. 2007; Bolle, Tan, and Zizzo 2014; Bosman and Van Winden 2002; Harth and Rieger 2017; Hopfensitz and Reuben 2009; Offerman 2002; Reuben and Van Winden 2008, 2010; Sanfey et al. 2003; Van Leeuwen et al. 2018; Xiao and Houser 2005) in contexts other than the Gift Exchange paradigm. Kirchsteiger, Rigo, and Rustichini (2006) observe a similar relationship in the Gift Exchange setting. Psychological research has also highlighted the role of anger in aggression and retaliation more generally, and that anger differs among individuals facing the same situation, that is, it is influenced by one’s personality traits as well as norms (Averill 1983; Eisenberger et al. 2004; Frijda 1986). Within economics, Battigalli, Dufwenberg, and Smith (2015) have recently modeled anger and aggression, using the tools of psychological game theory. Unlike Akerlof and Yellen (1990), they consider the possibility that anger subsides over time and that the desire for retaliation wanes.

Within economics, Battigalli, Dufwenberg, and Smith (2015) have recently modeled anger and aggression, using the tools of psychological game theory. Unlike Akerlof and Yellen (1990, 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.

5. Maximiano, Sloof, and Sonnemans (2007) show that the reciprocity found in Gift Exchange laboratory experiments is also observed when 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 et al. (2001); Charness and Kuhn (2007); Abele et al. (2010); Gächter and Thöni (2010); Cohn, Fehr, Herrmann, and Schneider (2014); and Charness et al. (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 both of their workers.

6. The reminder creates roughly the typical level of information a worker has when one starts a job (though the wage may have been set considerably in advance, workers typically have not forgotten how much they are getting paid when work begins on a project).

7. That is, 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. 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.
some suggestive evidence that anger plays a role in the behavior that we observe.

We obtain two main results in this study. First and foremost, the strength of the reciprocal wage-effort relationship is weaker after several weeks than it is within 3 hours. In particular, workers receiving low wages choose greater effort after the long delay, indicating that workers are more inclined to punish stingy employers in the short-term than later on. Second, we fail to find evidence in favor of the Fair Wage-Effort hypothesis. For short-term workers, we observe that reciprocity is not significantly stronger for wages below self-reported fair wages than for those above. There is a similar lack of asymmetry in the effect of wages below and above fair levels for long-term workers.

The rest of this paper is organized as follows: Section II describes the experimental design, Section III presents the hypotheses, Section IV reports the results, and Section V offers our conclusions.

II. EXPERIMENTAL DESIGN

In this section, we describe the structure of each of the three experiments. The first two experiments were conducted in the Netherlands and the third one in the United States. The first one is called Tilburg-L (Tilburg Low Efficiency), the second is called Tilburg-H (Tilburg High Efficiency), and the third is called Tucson-H (Tucson High Efficiency). Section II.A describes the procedures that were identical in the three experiments. Section II.B then covers the aspects that differed among them. It explains why three experiments were conducted and how the last experiment, Tucson-H, allows testing of some hypotheses beyond those evaluated in the first two experiments.

A. Procedures Common to all Experiments

Participants are assigned randomly to one of three roles: (1) Employer, (2) Short-term worker, or (3) Long-term worker. The roles are private information. Groups are formed, each consisting of one employer and two workers, one of each type. We describe the two workers as short- or long-term, based on when they make their effort decision. The participants are informed that the groups are randomly constituted and anonymous. 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 short-term worker must submit an effort level, $e_t$, within 3 hours after the end of the session. The long-term worker chooses an effort level $e_D$ within a 5-day interval beginning 4 weeks after the session. The workers, therefore, do not choose their effort in the laboratory. They send their choice to an email address provided to them. Participants are informed during the session that the long-term workers would receive an email reminder, containing the wage they have been awarded and a copy of the instructions. 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.

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 long-term worker she is matched with.

The payoffs to the three types of participants are given by:

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

Short-term worker: $d + w - c \times e_t$

8. The time intervals for Tilburg-L differ slightly. Short-term workers were instructed to send their effort choice within 2 hours, and long-term workers during a 24-hour window that began 1 month after the session. This was not enough time for some workers to send in their effort choice. We accepted effort choices from a few short-term workers after the 2-hour deadline. We informed long-term workers that they would have additional days to choose their effort, and accepted choices up to 8 days after the original deadline.

9. The instructions are provided in Appendix D and the email reminders are given in Appendix E.

10. Paying back the wage to employers when workers do not choose an effort level 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.

11. The payoff structure was presented to subjects in terms of these formulas, with the actual values in effect, rather than the variables indicated. See Charness, Fréchette, and Kagel (2004) for a discussion of the effects of different formats of presentation.
### TABLE 1
Procedural Details

| Parameters, Range of Variables, and other Details | Full Sample | 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     |           |
| $e$                                              | $[0, w]$    | $[0, 2.4]$| $[0, 4]$  |           |
| $w$                                              | $[0, 4 EUR]$| $[0 EUR, 2 EUR]$| $[0 USD, 4 USD]$| |
| Show-up fee                                      | 0 EUR       | 0 EUR     | 8 USD     |           |
| Average earnings employer                        | 10 EUR      | 8 EUR     | 22 USD    |           |
| Average earnings short-term worker               | 6 EUR       | 6 EUR     | 14 USD    |           |
| Average earnings long-term worker                | 6 EUR       | 6 EUR     | 14 USD    |           |
| Payment method                                   | Bank Transfer$^a$ | Bank Transfer | Cash and Mail | |
| Time period conducted                            | March–May 2015 | April–June, Nov.–Dec. 2015 | Jan.–March 2016 | |
| Location                                         | CentER Lab (Tilburg, The Netherlands) | CentER Lab (Tilburg, The Netherlands) | Economic Science Lab (Tucson, United States) |
| Number of participants                           | 431         | 138       | 99        | 194$^b$  |
| Attrition short-term workers$^c$                 | 4%          | 0%        | 3%        | 8%       |
| Attrition long-term workers                      | 10%         | 9%        | 9%        | 13%      |
| Net number participants                          | 410         | 134       | 95        | 181      |
| Net number employers                             | 144         | 46        | 33        | 65       |
| Net number short-term workers                    | 138         | 46        | 32        | 60       |
| Net number long-term workers                     | 128         | 42        | 30        | 56       |

$^a$Bank transfers are the most common method of payment in the Netherlands, even for small payments.

$^b$One subject in the role of long-term worker participated twice. We removed the second observation.

$^c$The wages are similar (i) for short- and long-term workers who do not respond, and (ii) for workers who do not respond and workers who do respond. This suggests that attrition is unlikely to drive differences between the behavior of short- and long-term workers. First, as explained in Section IV, we normalize wages by dividing the wage in each experiment by the maximum feasible wage offer in this experiment. The mean normalized wages of short- and long-term workers who did not respond are 0.54 ($SD = 0.31$, $N = 6$) and 0.54 ($SD = 0.32$, $N = 15$), slightly above half of the maximum possible wage and close to the mean normalized wage of 0.60 ($SD = 0.30$, $N = 266$) for workers who did respond. Second, we find no significant relationships between attrition and wage offers for either short-term or long-term workers using an ordinary least squares regression.

Long-term worker: $d + w - c \times e_D$

The ranges of possible wages, effort levels, effort costs, and output differ among the three experiments. The values are given in Table 1 along with other procedural details. 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 9 or 12 participants and lasted for 40 minutes. Participants are 51.5% male and their mean age is 22.1 ($SD = 3.0$).

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.$^{12}$

### B. Procedures Specific to each Experiment

**The Tilburg-L and Tilburg-H Experiments.** The procedures for the Tilburg-L experiment are described fully in Section II.A. The Tilburg-H experiment differs from Tilburg-L in one main aspect: it reduces the potential censoring of effort from above for low wages in Tilburg-L. That is, unlike in Tilburg-L and as in standard Gift Exchange experiments, the effort choice of workers in Tilburg-H is not limited by the wage

---

12. The game was programmed using the z-Tree software (Fischbacher 2007).
offered. In Tilburg–L, workers with low wages are restricted to choose low wages both in the short and long term. This means that we cannot observe a possible effect of the time delay on the effort choices of workers with low wages in Tilburg-L. Moreover, the parameters in effect in Tilburg-H also ensure greater productivity of effort. This is captured in the higher $b/c$ ratio. We increased the ratio to reduce censoring from below, in order to be better able to observe possible changes in effort over time.13 At the same time that he chooses a wage, the employer also requests the same nonbinding effort level from each worker. The effort request is included in the reminder to long-term workers.

The Tucson-H Experiment. This experiment extends the design of Tilburg-H to allow us to consider reciprocity in greater detail. Although the key ratio $b/c$ remains the same, there are other differences in the parameters, as indicated in Table 1. Moreover, workers are asked to state what they think would be a fair wage before learning their actual wage. It is emphasized to them that this fair wage would not be revealed to any other participant, including their employer, and therefore could not affect the wage they receive. Participants receive a show-up fee of 8 USD at the time of the session on top of their earnings from the game, which are paid at a later time.

III. HYPOTHESES

The first two hypotheses apply to all three experiments, and concern the effect of long versus short-time delays on effort choices. They originate from the results of prior experimental work. Hypotheses 3 and 4 apply exclusively to our third experiment, Tucson-H. They emerged after the first two experiments and guided the design of Tucson-H. These hypotheses posit asymmetries in worker behavior between cases in which the wage is perceived as fair and as unfair.

We first consider the relationship between effort and time delay, conditional on wage. That is, we test whether delay has any effect on the level of effort or on the sensitivity of effort to wages received. In the regression specification that we employ, these correspond to the intercept and slope coefficients, respectively. We allow for (1) long-term workers choosing a different effort level than short-term workers irrespective of the wage, and/or (2) long- and short-term workers exhibiting different sensitivities of effort to wage. Prior evidence that the level of effort may be affected by time delays comes from the dictator game literature (Dreber et al. 2016; Kovarik 2009), investigating the relationship between altruism and time lags. They find that when dictators have to split money that is payable in the future, they keep more for themselves. However, our environment is very different, and it is unclear whether effort levels would differ at the two time lags, controlling for wage level. Furthermore, as we discussed in the introduction, while the current literature suggests that the wage-effort relationship is weaker in the long term, the evidence is not extensive. Therefore, we take a conservative approach and propose the following null hypothesis.

**HYPOTHESIS 1:** For a given wage, short-term workers exert the same effort as long-term workers ($e_{lw} = e_{Dw}$).

Hypothesis 2 specifically considers whether a positive relationship between wage and effort appears with similar force within 3 hours and 1 month after workers learn their wages. As discussed in the introduction, the overall evidence from field experiments supports the contention that there is a decrease in reciprocity after a long delay. The Gift Exchange version of the Efficiency Wage Hypothesis (Akerlof 1982), on which most of the reciprocity labor literature is based, does not consider whether time should affect reciprocal behavior. In laboratory studies comparing reciprocity under immediate response and a short-time delay in other types of interaction, the results are mixed, but generally find a decrease in reciprocity. Bosman, Sonnemans, and Zeelenberg (2001) find that anger is associated with rejections in the ultimatum game and that a 1-hour delay does not decrease rejections. However, in their study, only a handful of participants reject offers so that a decrease would be difficult to detect. More recently, Grimm and Mengel (2011) find that a 10-minute delay reduces rejections in the ultimatum game, and Neo et al. (2013) observe a reduction of rejections in the ultimatum game with a similar delay, but no effect of delay in the investment game. Also related is the Internet-based experiment of Oechssler, Roider, and Schmitz (2015), who find that an opportunity to revise one’s decision 24 hours after having

13. Kessler (2013) found that a relatively high ratio tends to reduce the censoring of the effort data at the lower bound.
TABLE 2
Summary of Participants’ Average Decisions

| Experiment                              | Full Sample | N  | Tilburg-L | N  | Tilburg-H | N  | Tucson-H | N  |
|-----------------------------------------|-------------|----|-----------|----|-----------|----|-----------|----|
| Wage                                    | 0.600 (0.303) | 144 | 2.067 (1.288) | 46 | 1.218 (0.683) | 33 | 2.617 (1.014) | 65 |
| Effort of short-term worker             | 0.232 (0.285) | 138 | 0.343 (0.595) | 46 | 0.536 (0.727) | 32 | 1.395 (1.217) | 60 |
| Effort of long-term worker              | 0.274 (0.331) | 128 | 0.343 (0.596) | 42 | 0.610 (0.833) | 30 | 1.704 (1.397) | 56 |
| Fair wage indicated by short-term worker|             |     |           |    |           |    |           |    |
| Fair wage indicated by long-term worker |             |     |           |    |           |    |           |    |

Notes: Effort and wage range from 0 to 1 in the full sample, and their range differs between individual experiments (see Table 1 for details). In Tilburg-L, effort levels are limited above at the wage received. Standard deviations are in parentheses.

made a decision to accept or reject an offer in an ultimatum game decreases rejection in only one of their two conditions. Thus, our second null hypothesis is as follows.14

HYPOTHESIS 2: The sensitivity of effort to wage is the same for short- and long-term workers (de_L/dw = de_H/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, Fehr, and Goette (2015) are, to our knowledge, the only authors who test whether there is an asymmetric response about the self-reported fair wage. Because the study differs from ours, notably in terms of one-shot versus repeated effort decisions and of the timeframe employed, we maintain as our null hypothesis that the wage-effort relationship is the same for workers whose wage is below versus above their belief about what constitutes a fair wage.15

HYPOTHESIS 3: Short-term workers receiving a wage below their self-reported fair wage exhibit the same relationship between wage and effort as those receiving a wage above their fair level (de_L/dw_L = de_H/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 versus above the fair wage for short-term workers continues to exist for long-term workers. It is possible that negative reciprocity dissipates over time while 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 decrease over time. In the absence of any prior evidence, our null hypothesis is that any asymmetry in reciprocity is unaffected by a longer delay.

HYPOTHESIS 4: Short-term and long-term workers exhibit the same relationship between wage and effort if (a) the wage they receive exceeds their reported fair wage, and if (b) the wage they receive is less than their reported fair wage. (That is, (a) de_L/dw = de_H/dw for w > w_F, and (b) de_L/dw = de_H/dw for w < w_F).

IV. RESULTS

This section is organized in two parts. The first part presents an overview of the data, and the second reports the tests of our hypotheses.

A. Summary of the Data

In Table 2, we summarize the average values of participants’ decision variables for the full sample, as well as for the three individual experiments. We normalize wages and efforts in our full sample by dividing wages and efforts by their maximum possible values in each experiment.16 For the full sample and each individual experiment, the average effort of short- and long-term workers is close to the midpoint of the range of possible wage levels, which is well in excess of the minimum of zero. There are no significant differences in the short- and long-term average

14. Logically, Hypothesis 2 is nested in Hypothesis 1, which asserts that effort, conditional on wage, is identical for short-term and long-term workers. However, Hypothesis 2 is of special interest because it captures any reciprocal behavior of the workers.

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

16. We normalize the data by dividing wages and efforts by 4 in Tilburg-L and Tucson-H, and by 2 for short and efforts by 2.4 in Tilburg-H.
TABLE 3
Average Effort by Wage Range, for Short- and Long-Term Workers

| Wage Range | Short-Term Workers | Long-term Workers |
|------------|--------------------|-------------------|
|            | Average Effort     | N                  | Average Effort | N                  |
| Full sample|                    |                    |                |
| 0.00–0.50  | 0.085 (0.149)      | 53                 | 0.211 (0.286)  | 50                 |
| 0.51–1.00  | 0.324 (0.310)      | 85                 | 0.314 (0.353)  | 78                 |
| Tilburg-L  | 0.0–2.0            | 0.175 (0.439)      | 24              | 0.362 (0.571)      | 21                 |
|            | 2.1–4.0            | 0.527 (0.694)      | 22              | 0.324 (0.634)      | 21                 |
| Tilburg-H  | 0.0–1.0            | 0.145 (0.157)      | 11              | 0.782 (1.001)      | 11                 |
|            | 1.1–2.0            | 0.743 (0.823)      | 21              | 0.511 (0.730)      | 19                 |
| Tucson-H   | 0.0–2.0            | 0.622 (0.812)      | 18              | 1.128 (1.114)      | 18                 |
|            | 2.1–4.0            | 1.726 (1.218)      | 42              | 1.976 (1.447)      | 38                 |

Notes: The numbers presented are averages. For the full sample, wage and effort are normalized so that both range from 0 to 1. For the individual experiments, the variables are not normalized. In Tilburg-L and Tucson-H, wage and effort range from 0 to 4, but effort is limited above by the wage in Tilburg-L. In Tilburg-H, wage ranges from 0 to 2 and effort from 0 to 2.4. The standard deviations are in parentheses.

In Table 3, we present the average effort levels for each half of the wage range, for both types of workers. In the full sample, for short-term workers, an increasing relationship between wage and short-term effort appears. For long-term workers, the wage-effort relationship seems to be weaker. Specifically, wages in the lower half of the wage range lead to greater effort for long- than for short-term workers. Moreover, the picture is similar for each individual experiment. This tendency for greater long-term effort in the lower half 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 low wages to select high effort levels (e.g., workers with a wage of 0 can only choose an effort of 0). We can make similar observations from Figure 1, which illustrates the relationship between the wages workers receive and the effort they choose. The graphs contain the normalized data from the full sample, for short- and long-term workers separately. We highlight, in pale orange rectangles, the effort at wages 0 and 0.1 in Tilburg-L, which are bounded above at 0 and 0.1, respectively, in both the short and long term. We include the fitted regression lines in the figure.

B. Evaluation of Hypotheses

Hypotheses 1 and 2. The first two hypotheses concern differences in effort between short- and long-term workers. To test for differences, we pool the normalized data of workers in our three experiments and estimate the following Tobit specification.

\[ e_j = \alpha + \beta \times wage_j + \theta \times wage_j \times D_j + \alpha_{TilbH} \times TilbH_j + \alpha_{TucH} \times TucH_j + \gamma \times D_j + \epsilon_j \]

In this equation, \( e_j \) denotes the effort of worker \( j \), \( wage_j \) is the wage \( j \) has received, and \( D_j \) indicates a long delay (equals 0 for short-term workers and 1 for long-term workers). To simplify, we will usually write “delay” instead of “long delay.” There is also an interaction term between wage and the delay included in the specification. We include two dummies, for the Tilburg-H and Tucson-H experiments. The estimates are presented in column 1 of Table 4. In column 2, we include interaction terms between the wage and Tilburg-H and the wage and Tucson-H. In column 3, we add additional interaction terms between the wage and delay that are specific to Tilburg-H and Tucson-H.

The first hypothesis is that the time delay has no effect on effort for a given wage. That is, in terms of a regression of effort on wage, the delay affects neither the intercept (effort for the minimum wage) nor the slope (the wage-effort sensitivity). We assess this proposition by testing whether the coefficients of \( Wage \times Delay \) and Delay in column 1 are both equal to zero. Note that our first hypothesis differs from claiming that the average effort is the same for both types of

17. Appendix A shows the same table for each quarter of the wage range; the pattern of effort is similar.

18. We use a Tobit specification because there is a fair amount of lower censoring (35% of observations) for the effort variable at 0, and some censoring above (6%) at the maximum effort choice, as can be seen in Figure 1. We assume normalized effort to be censored below 0 and above 1 for simplicity in analyzing the data, even though in Tilburg-L effort is actually censored at the wage offered. We use robust standard errors in all regressions.
worker (as reported in the previous subsection, the average effort does not differ). This is because similar average effort can hide a pattern in which a long-time delay causes effort to change in opposite directions for workers who receive low wages and workers who receive high wages. We reject the hypothesis (F-test, p value is .04). We obtain the same result for the coefficients of columns 2 or 3. This yields our first result.

RESULT 1. For a given wage, short- and long-term workers make significantly different effort choices.

We now consider the nature of the effect of the time delay on the behavior of short- and long-term workers. Our second hypothesis is that the wage-effort relationship does not change with delay. The coefficient of Wage is positive and significant (columns 1–3, p values are <.001), replicating the common finding that higher wages lead to higher effort on the part of workers.\textsuperscript{19} The coefficient of Wage × Delay in column 1 is, however, negative and significant (p value is .02). As column 2 shows, this is robust to including wage terms for Tilburg-L and Tucson-H in the regression. This indicates that the long-time delay decreases the marginal effect of the wage on effort provided. Moreover, we do not find significant reciprocity in the long term (for the F-tests of the restriction that Wage + Wage ×

\textsuperscript{19} Though not shown here, short-term workers providing their effort earlier versus later within the permitted time interval exhibit the same wage-effort relationship. That is, short-term workers responding within 15 or 30 minutes after the session show no differences in their wage-effort relationship compared to short-term workers responding later in the 3-hour interval. This suggests that they recall the wage they receive equally well later in the 3 hour interval as they do in the first few minutes. Because the long-term workers are reminded of their wage at the beginning of the window in which they make their decision, we take this as confirmatory evidence that the information that short- and long-term workers have at the time they choose their effort is similar.
TABLE 4
Estimates of Tobit Regression (1)

|            | (1) Effort | (2) Effort | (3) Effort |
|------------|------------|------------|------------|
| Wage       | 0.680***   | 0.672***** | 0.690****  |
|            | (0.121)    | (0.169)    | (0.171)    |
| Wage × D   | −0.478**   | −0.480**   | −0.525**   |
|            | (0.209)    | (0.200)    | (0.223)    |
| D          | 0.355**    | 0.358***   | 0.349***   |
|            | (0.138)    | (0.130)    | (0.129)    |
| Wage × TucH| −0.287     | −0.247     | −0.226     |
|            | (0.242)    | (0.254)    | (0.226)    |
| Wage × TilbH| 0.305     | 0.235      | 0.228      |
|            | (0.218)    | (0.208)    | (0.177)    |
| Wage × D × TilbH| −0.082 |
|            |           | (0.208)    |
| Wage × D × TucH| 0.161 |
|            |           | (0.177)    |
| TilbH      | 0.229***   | 0.410**    | 0.408**    |
|            | (0.074)    | (0.167)    | (0.166)    |
| TucH       | 0.416***   | 0.211      | 0.207      |
|            | (0.059)    | (0.139)    | (0.137)    |
| Constant   | −0.529**** | −0.522**** | −0.514**** |
|            | (0.093)    | (0.115)    | (0.114)    |
| Prob. > F  | 0.000      | 0.000      | 0.000      |
| Pseudo R²  | 0.200      | 0.218      | 0.223      |
| N          | 266        | 266        | 266        |

Note: N denotes the sample size, which includes both types of workers, split almost evenly.
* p = .10; ** p = .05; *** p = .01; **** p = .001.

TABLE 5
Spearman Correlation between Wage and Effort

|                     | Full Sample | N   |
|---------------------|-------------|-----|
| Short-term workers  | 0.467 [<.001] | 138 |
| Long-term workers   | 0.173 [0.051] | 128 |

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

RESULT 2. The wage-effort relationship is stronger within 3 hours, than 4 weeks after, workers learn their wage. Specifically, low wages generate less shirking in the long term.

Hypotheses 3 and 4. Our two remaining hypotheses apply only to our third experiment, Tucson-H, and concern the asymmetry in reciprocity about the fair wage. Our sample of workers consequently decreases from 266 to 116 for the analysis. There is therefore a relatively small number of workers falling in the categories that we compare to one another.

We analyze the behavior of workers as a function of their wage and an unfairness indicator variable I that denotes whether their wage is below their self-reported fair wage. The indicator takes value 1 if the wage is below fair levels, and 0 otherwise. In order to give an idea of the wage-effort relationship as well as the unfairness indicator-effort relationship, Table 6 shows the Spearman correlations between I and effort, separately for short- and long-term workers. We conduct Spearman’s rank correlation tests and present the p values in square brackets. We make two observations. First, the wage-effort correlation is positive and significant for both short- and long-term workers, although the magnitude of the short-term correlation is greater (the long-term correlation amounts to 37% of the short-term one).

We report, in Appendix A, that the wage-effort relationship also weakens over time if we use an approach based on wage ranges instead of the wages themselves. As the descriptive statistics suggest, we show that this decrease is driven by workers who receive wages in the lower half of the range of possible wages. That is, long-term workers offered low wages provide more effort than short-term workers offered the same low wages.20 This pattern provides the basis for our second result.

20. We note that the equitable payoff—which we take as an equal payoff to the employer and to each employee, once the employer has decided on a wage—is achieved here when the employer offers a standardized wage of approximately 0.5. This corresponds to offering half of the maximum possible wage, which is 1. This means that those short- and long-term workers who are offered a wage below the wage generating the equitable payoff are the ones choosing different effort levels. That is, long-term workers treated inequitably provide greater effort than short-term workers.
TABLE 6
Spearman Correlation between Wage and Effort as well as between the Unfairness Indicator I and Effort

| Wage | I | N |
|------|---|---|
| Short-term workers | 0.463 [<0.001] | −0.188 [0.151] | 60 |
| Long-term workers  | 0.298 [0.026]  | −0.210 [0.121] | 56 |

Notes: The unfairness indicator I takes value 1 if Fair Wage > Wage and value 0 otherwise. We indicate the two-sided p values from Spearman’s rank correlation tests in square brackets.

correlation corresponds to 64% of the short-term one). Second, for both short- and long-term workers, the unfairness indicator-effort correlation is negative, indicating that offering wages below the fair wage decreases effort, but the relationship is not statistically significant.

To take into account both the effect of the absolute wage and the effect of the wage being below the fair level, we estimate the following Tobit regression:

\[
(2) \quad e_j = \alpha + \beta_1 \times \text{wage}_j + \theta_1 \times \text{wage}_j \times D_j + \beta_2 I_j + \theta_2 I_j \times D_j + \gamma \times D_j + \epsilon_j
\]

The unfairness indicator variable \(I_j\) captures whether worker \(j\) receives a wage below the fair wage. We include a variable for the wage, and to detect differences between the short and the long term, we interact the wage and the indicator with the time delay \(D_j\).

21. Instead, we could analyze the behavior of workers as a function of their wage and whether their wage is below their reported fair wage by a sufficiently large amount. The cutoff that corresponds to the Fair Wage-Effort hypothesis is 0 USD, that is, any wage that is below fair levels. For the purpose of our analysis, choosing higher cutoffs might allow us to decrease the collinearity between the wage and the indicator variable for whether the wage is below the fair wage. However, the Pearson correlation between the wage and whether the wage is below fair levels remains near 0.65 for a wide range of cutoffs (i.e., 0.50, 1.00, 1.50, or 2.00 USD). Consequently, choosing cutoffs of 0.50, 1.00, 1.50, or 2.00 USD would provide the same Results 3 and 4 as the cutoff of 0 USD. Note that, among short-term workers, 42 receive a wage below the fair wage and 23 get a wage above, and among long-term workers, 44 have a wage below the fair wage and 20 have a wage above. Moreover, 26 short-term workers receive a wage at least 1 USD below the fair wage and 39 have a wage above, and 31 long-term workers have a wage at least 1 USD below the fair wage, while 33 get a wage above.

22. Note that we do not include a variable for the self-reported fair wage among the explanatory variables because it would be highly correlated with our unfairness indicator.

The estimates are presented in columns 1–3 of Table 7. Column 1 provides estimates from the regression including only the wage variables. Column 2 includes the two variables containing the unfairness indicator, but excludes the wage variables. Column 3 employs the full regression specification.

In column 1, we find a significant positive wage-effort relationship (\(Wage, p\) value is <.001), but, unlike for the full sample, the decrease in this relationship over time is not significant (\(Wage \times D, p\) value is .38). This difference can be explained by the loss in statistical power compared to the full sample. In column 2, the effect of a wage below the fair wage is negative but insignificant (\(I, p\) value is .16), and the attenuation of the effect is also not significant (\(I \times D, p\) value is .64).

To evaluate Hypothesis 3—that short-term workers would exhibit the same wage-effort relationship for wages below and above their self-reported fair wage—we use column 3, as it allows us to isolate the effect of the wage being below fair levels. We find that the wage has a significant impact (\(Wage, p\) value is <.001), but that having a wage below the fair wage in itself does

23. For instance, simulations show that if we draw 10,000 random subsamples of 116 workers (58 short-term and 58 long-term workers, without replacement) from our full sample, we detect a decrease in reciprocity in 27% of them at the 5% significance level, and in 42% of them at the 10% significance level.
not decrease effort significantly (the effect of $I$ in column 3 is actually positive, though only close to marginally significant, $p$ value is .102).\footnote{In view of the fact that negative reciprocity has been found to be stronger than positive in several other studies, using a one-sided $p$ value instead would also be reasonable here, and this would make the difference very close to significant at the 5% level.}

We refer again to column 3 in order to examine Hypothesis 4—that the asymmetry between negative and positive reciprocity would be the same for short- and long-term workers. We see that the signs of the two coefficients of the interactions with time delay suggest an attenuation in the effect of the wage, as well as the effect of the wage being below fair levels. The decrease in the effect of the wage is marginally significant ($Wage \times D$, $p$ value is .097), but the change over time in the effect of having a wage below what is considered fair is not significant ($I \times D$, $p$ value is .16). Results 3 and 4 describe our findings concerning Hypotheses 3 and 4.

RESULT 3. For short-term workers, the sensitivity of effort to wage is similar for wages below fair levels than for those above.

RESULT 4. The difference in the sensitivity of effort to wage between wages below and above fair levels does not differ significantly between short-term and long-term workers.

In Appendix B, we present an exploratory analysis regarding the relationships between workers’ wages, their emotions, and their subsequent effort choices, using data on emotional state obtained with Facereading software. Our results are suggestive of a negative relationship between workers’ anger and their effort levels in the short term, though not the long term.\footnote{In addition, Appendix C details employers’ expectations of short-term and long-term efforts choices.}

V. CONCLUSION

Akerlof (1982) formalizes the Gift-Exchange version of the Efficient Wage hypothesis, in which employers offer wages that are higher than the market-clearing wage and workers reciprocate higher wages with greater effort. While much support for the existence of reciprocal provision of effort by workers has been gathered in short-term interactions, studies of whether reciprocal behavior persists over longer time periods have generally found that it persists within a work day, but decreases afterwards (Bellemare and Shearer 2009; Cohn, Fehr, and Goette 2015; Gilchrist, Luca, and Malhotra 2016; Gneezy and List 2006; Hossain and List 2012; Kube, Maréchal, and Puppe 2012, 2013). Here, we consider whether the reciprocity workers exhibit in their effort choices differs under a brief versus a long delay (within 3 hours vs. after 4 weeks) from the time at which they learn their wage. We test whether a positive wage-effort relationship is observed in an environment that is more decontextualized (in the laboratory) and simpler (a one-shot interaction) than previous studies concerned with reciprocity over long time periods. This allows us to clearly isolate any decrease in reciprocal behavior engendered by the passing of time over a long period.

In our third, more extensive experiment, we also consider one aspect of the Fair Wage-Effort hypothesis of Akerlof and Yellen (1988, 1990) that might differ across timeframes: the contention that workers shirk in retaliation for wages below the level that they perceive as fair, but do not reward wages greater than the fair level with more effort. The existence of this asymmetry has been supported by the field experiment reported by Cohn, Fehr, and Goette (2015). We examine whether this asymmetry exists in our setting, both in the short and long term.

Our main finding is that the strength of the reciprocal wage-effort relationship observed for short-term workers (who choose their effort within 3 hours of learning the wage) decreases for long-term workers (who wait 1 month before choosing their effort). We also find that workers offered low wages are the ones exhibiting different behavior over time. That is, long-term workers with low wages choose greater effort than short-term workers with the same low wages.

We obtain two additional findings from our third experiment that fail to support the Fair Wage-Effort hypothesis. The first is that, once we take into account the effect of the wage, workers do not reciprocate significantly more strongly for a wage that falls below the self-reported fair wage. The second is that this failure of the Fair-Wage effort hypothesis is present under both the short and long delays. Because of the reduced sample size on which these last two findings rest, they should not be taken as absolutely definitive. The signs of the estimated coefficients point in the direction of stronger reciprocation for wages below fair levels, as well as an attenuation of this asymmetry for long term...
workers, but the effects are below conventional thresholds of significance. Future research can shed more light on the presence of these two relationships.

Overall, our findings add to the evidence gathered by previous studies suggesting that workers’ reciprocity weakens over long time periods. Moreover, we report that the magnitude of the wage appears to influence the effect of delays; wages need to be low enough for reciprocity to change over time. For employers, our results imply that examination of the timing of the task at hand relative to when wages are determined can benefit those trying to induce high effort from their employees.

REFERENCES

Abeler, J., S. Altmann, S. Kube, and M. Wibral. “Gift Exchange and workers’ Fairness Concerns: When Equality Is Unfair.” Journal of the European Economic Association, 8(6), 2010, 1299–324.

Akerlof, G. A. “Labor Contracts as Partial Gift Exchange.” Quarterly Journal of Economics, 97(4), 1982, 543–69.

Akerlof, G. A., and J. L. Yellen. “Fairness and unemploy-

Averill, J. R. “Studies on Anger and Aggression: Implications for Theories of Emotion.” American Psychologist, 38(11), 1983, 1145–60.

Battigalli, P., M. Dufwenberg, and A. Smith “Frustration and Anger in Games.” Working Paper, Bocconi University, 2015.

Bellemare, C., and B. Shearer. “Gift Giving and Worker Productivity: Evidence from a Firm-Level Experiment.” Games and Economic Behavior, 67(1), 2009, 233–44.

Ben-Shakhar, G., G. Bornstein, A. Hopfensitz, and F. van-Winden. “Reciprocity and Emotions in Bargaining Using Physiological and Self-Report Measures.” Journal of Personality and Social Psychology, 92(2), 2007, 314–23.

Bolle, F., J. H. Tan, and D. J. Zizzo. “Vendettas.” American Economic Journal: Microeconomics, 6(2), 2014, 93–130.

Bosman, R., J. Sonnemans, and M. Zeelenberg. “Emotions, Rejections, and Cooling off in the Ultimatum Game.” Working Paper, University of Amsterdam, 2001.

Bosman, R., and F. Van Winden. “Emotional Hazard in a Power-to-Take Experiment.” Economic Journal, 112(476), 2002, 147–69.

Brandts, J., and G. Charness. “Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games.” Experimental Economics, 2(3), 2000, 227–38.

Brown, M. A., A. Falk, and E. Fehr. “Relational Contracts and the Nature of Market Interactions.” Econometrica, 72(3), 2004, 747–80.

Casoria, F., and A. Riedl. “Experimental Labor Markets and Policy Considerations: Incomplete Contracts and Macroeconomic Aspects.” Journal of Economic Surveys, 27(3), 2013, 398–420.

Charness, G. “Attribution and Reciprocity in an Experimental Labor Market.” Journal of Labor Economics, 22(3), 2004, 665–88.

Charness, G., R. Cobréyes, J. A. Lacomba, F. Lagos, and J. M. Perez. “Social Comparisons in Wage Delegation: Experimental Evidence.” Experimental Economics, 19(2), 2016, 433–59.

Charness, G., G. R. Fréchette, and J. H. Kagel. “How Robust Is Laboratory Gift Exchange?” Experimental Economics, 7(2), 2004, 189–205.

Charness, G., and P. Kuhn. “Does Pay Inequality Affect Worker Effort? Experimental Evidence.” Journal of Labor Economics, 25(4), 2007, 693–723.

Cohn, A., E. Fehr, and L. Goette. “Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment.” Management Science, 61(8), 2015, 1777–94.

Cohn, A., E. Fehr, B. Herrmann, and F. Schneider. “Social Comparison and Effort Provision: Evidence from a Field Experiment.” Journal of the European Economic Association, 12(4), 2014, 877–98.

Cooper, D. J., and J. H. Kagel. “Other Regarding Preferences: A Selective Survey of Experimental Results.” in Handbook of Experimental Economics, 2, edited by J. H. Kagel and A. E. Roth. Princeton, NJ: Princeton University Press, 2016, 217–289.

Dufwenberg, M., and G. Kirchsteiger. “A Theory of Sequential Reciprocity.” Games and Economic Behavior, 47(2), 2004, 268–98.

Dreber, A., D. Fudenberg, D. K. Levine, and D. G. Rand. “Self-Control, Social Preferences and the Effect of Delayed Payments.” Working Paper, Washington University in Saint Louis, 2016.

Eisenberger, R., P. Lynch, J. Aselage, and S. Rohdieck. “Who Takes the Most Revenge? Individual Differences in Negative Reciprocity Norm Endorsement.” Personality and Social Psychology Bulletin, 30(6), 2004, 787–99.

Fehr, E., and A. Falk. “Wage Rigidity in a Competitive Incomplete Contract Market.” Journal of Political Economy, 107(1), 1999, 106–34.

Fehr, E., S. Gächter, and G. Kirchsteiger. “Reciprocity as a Contract Enforcement Device: Experimental Evidence.” Econometrica, 65(4), 1997, 833–60.

Fehr, E., L. Goette, and C. Zehnder. “A Behavioral Account of the Labor Market: The Role of Fairness Concerns.” Annual Review of Economics, 1(1), 2009, 355–84.

Fehr, E., E. Kirchler, A. Weichbold, and S. Gächter. “When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets.” Journal of Labor Economics, 16(2), 1998a, 324–51.

Fehr, E., G. Kirchsteiger, and A. Riedl. “Does Fairness Prevent Market Clearing? An Experimental Investigation.” Quarterly Journal of Economics, 108(2), 1993, 437–59.

Fehr, E., G. Kirchsteiger, and A. Riedl. “Do Involuntary Unemployment and Non-compensating Wage Differentials in an Experimental Labour Market?” Economic Journal, 106(434), 1996, 106–21.

Fischbacher, U. “Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments.” European Economic Review, 42(1), 1998b, 1–34.

Frijda, N. H. The Emotions. Cambridge, MA: Cambridge University Press, 1986.

Gächter, S., and A. Falk. “Reputation and Reciprocity: Consequences for the Labour Relation.” Scandinavian Journal of Economics, 104(1), 2002, 1–26.

Gächter, S., and C. Thöni. “Social Comparison and Performance: Experimental Evidence on the Fair
Wage–Effort Hypothesis.” *Journal of Economic Behavior & Organization*, 76(3), 2010, 531–43.

Gilchrist, D. S., M. Luca, and D. Malhotra. “When 3 + 1 > 4: Gift Structure and Reciprocity in the Field.” *Management Science*, 62(9), 2016, 2639–50.

Gneezy, U., and J. A. List. “Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments.” *Econometrica*, 74(5), 2006, 1365–84.

Grimm, V., and F. Mengel. “Let me Sleep on it: Delay Reduces Rejection Rates in Ultimatum Games.” *Economics Letters*, 111(2), 2011, 113–15.

Guth, W., M. Königstein, J. Kovács, and E. Zala-Mező. “Fairness within Firms: The Case of One Principal and Multiple Agents.” *Schmalenbach Business Review: ZFBF*, 53(2), 2001, 82–101.

Harth, N. S., and T. Regner. “The Spiral of Distrust: (Non-)Cooperation in a Repeated Trust Game Is Predicted by Anger and Individual Differences in Negative Reciprocity Orientation.” *International Journal of Psychology*, 52, 2017, 18–25.

Hannan, R. L., J. H. Kagel, and D. V. Moser. “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), 2002, 923–51.

Hennig-Schmidt, H., A. Sadrieh, and B. Rockenbach. “In Search of workers’ Real Effort Reciprocity—A Field and a Laboratory Experiment.” *Journal of the European Economic Association*, 8(4), 2010, 817–37.

Oechssler, J., A. Roeder, and P. W. Schmitz. “Cooling off in Negotiations: Does It Work?” *Journal of Institutional and Theoretical Economics*, 171(4), 2015, 565–88.

Offerman, T. “Hurting Hurts More than Helping Helps.” *European Economic Review*, 46(8), 2002, 1423–37.

Hopfensitz, A., and E. Reuben. “The Importance of Emotions for the Effectiveness of Social Punishment.” *Economic Journal*, 119(540), 2009, 1534–59.

Hossain, T., and J. A. List. “The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations.” *Management Science*, 58(12), 2012, 2151–67.

Kessler, J. B. “When Will there Be Gift Exchange? Addressing the Lab-Field Debate with Laboratory Gift Exchange Experiments.” Working Paper, University of Pennsylvania, 2013.

Kirksteiger, G. L., Rigotti, and A. Rustichini. “Your Morals Might Be your Moods.” *Journal of Economic Behavior & Organization*, 59(2), 2006, 155–72.

Kovarik, J. “Giving it Now or Later: Altruism and Discounting.” *Economics Letters*, 102(3), 2009, 152–4.