



# Managerial Payoff and Gift-Exchange in the Field

Florian Englmaier<sup>1</sup> · Stephen Leider<sup>2</sup>

Published online: 18 November 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

We conduct a field experiment where we vary both the presence of a gift-exchange wage and the effect of the worker’s effort on the manager’s payoff. Results indicate a strong complementarity between the initial wage-gift and the agent’s ability to “repay the gift”. We control for differences in ability and reciprocal inclination and show that gift-exchange is more effective with more reciprocal agents. We present a principal-agent model with reciprocal subjects that motivates our findings. Our results help to reconcile the conflicting evidence on the efficacy of gift-exchange outside the lab.

**Keywords** Incentives · Field experiments · Gift-exchange · Reciprocity

## 1 Introduction

Firms must regularly address the problem of providing incentives to employees when actions are not at all or hardly contractible. The standard approach in economics has been to focus on analyzing the optimal explicit incentive schemes: tying the level of the worker’s compensation to the amount of output produced, which serves as a (noisy) measure of employee effort.<sup>1</sup> However, the importance of fairness and social preferences—especially for the work relationship—has long been documented. Starting with Akerlof (1982), a literature has developed that considers gift-exchange as an alternative source of incentives in the workplace. According

---

<sup>1</sup> An extensive theoretical literature—e.g., Holmström (1979) or Grossman and Hart (1983)—has emphasized the importance of strong monetary incentives to induce effort.

---

✉ Florian Englmaier  
florian.englmaier@econ.lmu.de

Stephen Leider  
leider@umich.edu

<sup>1</sup> Department of Economics and Organizations Research Group, LMU Munich, Kaulbachstrasse 45, 80539 Munich, Germany

<sup>2</sup> University of Michigan, Ross School of Business, Technology & Operations Department, 701 Tappan Street R4424, Ann Arbor, MI 48109, USA

to this gift-exchange theory, firms can induce above-minimal effort from agents—even in the absence of explicit pay-for-performance incentives—since such-inclined agents may reciprocate generous wage payments with higher effort exertion: a positive wage-effort relationship takes effect as workers repay the initial “gift” of a generous wage with a “return gift” of above-minimum effort.

Recently there have been conflicting results on the significance of gift-exchange as a motivating force outside of the lab. While some papers have found a robust pattern of gift-exchange in field settings, others have failed to find the positive relationship of initial and return gift at all or have documented only effects that persist for very short periods; cf. DellaVigna et al. (2016). The varying effectiveness of gift-exchange in different settings suggests that the efficacy of gift-exchange incentives depends on details of the environment. We add to this research by providing results that suggest an avenue to reconcile those findings: In our field experiment we do not find evidence for an overall positive effort response merely from an initial wage gift. However, the gift’s efficacy is substantially improved if the manager benefits more strongly from a worker’s high effort. That is, gift-exchange can induce effort if workers are able to repay the gift to the manager. Moreover, we document that gift-exchange works more effectively with subjects that we classify as reciprocal via a personality test,<sup>2</sup> and that the efficacy of gift-exchange does not dissipate over the course of the experiment. We conclude from our results that while gift-exchange may not be effective as an incentive in all settings, it can be a powerful incentive device in the proper job context: such as in our setting when managers have performance-related incentives, and when it is directed to the right employees, who are most likely to be reciprocal.

For our field experiment we hired temporary workers for a data entry job.<sup>3</sup> The workers entered historical data from the 1849 Prussian Census. In total we had 59 workers entering data during a five hour shift in the Harvard Business School computer lab where the data entry took place. The job was advertised to the worker by a temp worker agency at the standard hourly wage of \$13; however for 30 of the workers we increased the hourly wage to \$18 upon their arrival. We explained to all of the workers that we were hired by two professors to organize the entry of these data. For 15 workers in each of the \$13 and the \$18 groups we emphasized the importance of them working hard for us by explaining that we would receive a bonus of 50% if the job was done ‘by the end of the week’ (*Bonus* treatment). For the ‘control’ groups in both of the \$13 and the \$18 conditions we did not inform the workers of this bonus (*No Bonus* treatment). We also asked the workers to fill out a short version of the “Big 5” personality test, to give us a measure of non-cognitive skills.<sup>4</sup>

<sup>2</sup> We were, in 2007, among the first to take seriously the documented heterogeneity in the population along the reciprocity dimension and to test a key tenet of reciprocity models: that treatment effects should be brought about by reciprocal subjects.

<sup>3</sup> On a more methodological level we think this paper was, at the time of writing in 2007, the first to recruit subjects for a natural field experiment via a temp worker agency and to exploit additional data that could be garnered from this specific subject pool.

<sup>4</sup> The “Big 5” personality factors capture five distinct aspects of an individual’s personality: extraversion; agreeableness; conscientiousness; emotional stability; and imagination. These factors are typically measured by having subjects agree or disagree with an extensive series of statements.

In line with previous field experiments, we do not find an effort increase in response to the higher wage in the *No Bonus* treatment. However, for the workers in the *Bonus* treatment, a higher wage leads to a significant increase in worker output. Hence, there is a strong complementarity between the wage gift and the resulting payoff for the manager.<sup>5</sup> Furthermore, when we separate workers based on their *agreeableness* (a personality trait associated with standard lab measures of reciprocity), we find that the positive effect of high wages on effort is driven entirely by highly agreeable (strong positive reciprocity) workers, with low reciprocity workers showing either a zero or a negative response to a higher wage. We consider the finding that the strength of gift-exchange is positively correlated with measures of reciprocal inclination as absolutely necessary to lend credibility to gift-exchange based explanations of motivation in the labor market. Though apparently obvious, to the best of our knowledge this study was, at the time of writing it in 2007/08, the first to document this fact in the field.

We also examine the strength of the gift-exchange response over time. In contrast to other studies, e.g. Gneezy and List (2006), we find no weakening of positive responses to wage gifts over time. To the contrary, any negative effects of the wage gift disappear in the later stages of the task, and we find an overall strongly positive effect of our treatment manipulations in the second half of the experiment.

We use two different measures of effort in our analysis: gross data entered, and an error-corrected measure of data entered. The estimates that were derived from these two measures are qualitatively and quantitatively very similar, which suggests that any effort responses work along the quantity margin only; the responses leave the quality margin of effort unaffected.<sup>6</sup>

We can rationalize these results with an agency model that captures reciprocal preferences, where we show that there is complementarity between gift and ability to repay the gift. In the model a risk-neutral firm hires a risk-averse worker to exert non-contractible effort. The novel feature of the model is that the worker is reciprocal: the worker's utility increases in the principal's profit whenever the worker receives a rent in excess of his outside option. Thus, when the firm is generous to the worker by giving him additional compensation, the worker desires to provide in turn something of value to the firm. The worker's reciprocal attitude can now be used by the firm to align the worker's preferences with those of the firm, which thus generates intrinsic motivation. The comparative statics show that *ceteris paribus* the worker's optimal effort choice increases in the initial wage gift and his ability to repay the gift. The corresponding cross derivative is also positive, which indicates the complementarity of the instruments.

An extensive body of evidence has developed that demonstrates reciprocal behavior and gift-exchange in laboratory experiments. Fehr and Gächter (2000) summarize results from earlier studies and highlight several key results: (i) Average wages in the experiments are above the minimal wages and leave workers with rents; (ii)

<sup>5</sup> This result is akin to the findings that are summarized in Ichniowski and Shaw (2003) on the complementarities among various HR policy instruments.

<sup>6</sup> This is in contrast to Kim and Slonim (2009), who find gift-exchange mainly along the quality margin.

There is a positive wage-effort relationship; and (iii) These results are robust to various institutions, to competition, and to high stakes.<sup>7</sup> The laboratory experiment by Hennig-Schmidt et al. (2010) is closely related to our paper. Hennig-Schmidt et al. (2010) present a real-effort laboratory experiment and show that a positive wage-effort relation as implied by gift-exchange prevails only if information on the manager's surplus is provided to the experimental workers. This indicates—as is predicted by our model—that the manager's surplus is an important determinant of the effectiveness of gift-exchange relations. Note, however, that Hennig-Schmidt et al. (2010) do not vary the surplus accruing to the manager nor do they collect the additional information necessary to test our hypotheses.

It is beyond the scope of this paper to give a comprehensive overview of the vast set of excellent papers on gift-exchange that have been written in the last decade. See DellaVigna et al. (2016) or Esteves-Sorenson (2018) for excellent overviews of this literature. These two papers also address the controversial discussion about the validity of the lab results for gift-exchange in the field. We take the mixed findings of these field studies as evidence that the efficacy of gift-exchange depends on subtle details of the field situation. While Falk (2007) finds strong evidence for gift-exchange in a field experiment with charitable donations, Gneezy and List (2006) argue that the effect of gift-exchange in the field is only minor, fast disappearing, and overall not a viable employment strategy. In Gneezy and List, students are hired for a day job in a library, and half of them get a surprise rise of their hourly pay. Gneezy and List document that, other than in our field experiment, there is only a short lived effect of this gift on the students' effort. Overall the 'firm' would have fared better hiring more students for the lower wage rate. Kube et al. (2013) replicate the Gneezy and List study and also find no effect of a wage gift but document a strong negative effect in response to a wage cut. In a comparable design, Kube et al. (2012) document a strong positive effect of non-monetary gifts, such as wrapped thermos bottles, on students' effort. Note that in neither of these cases were the subjects given any indication that the manager who provided the higher wage would benefit directly from increased productivity and the performed jobs were not ones where an employee would expect such a compensation structure. Becker et al. (2012) study a field experiment where a random sub-sample of participants in the Swiss Labor Force Survey received vouchers for training courses. The authors find evidence for long-term (six months) gift-exchange, as voucher recipients are more likely to participate in future survey waves. As the authors can track actual voucher redemption, they are able to document that this long-lasting gift-exchange relationship is most pronounced for the sub-group that had redeemed their vouchers.

In light of the above studies, the current study makes two contributions: On the one hand, as we also elicit measures of ability and personality traits, we were in 2007/2008 (to the best of our knowledge) the first to document the heterogeneity of the effect of gift-exchange across different types of employees. On the other hand, we add to the two "standard" features of a gift-exchange experiment—a

---

<sup>7</sup> For further reference see also Fehr and Gächter (2000) or Fehr and Schmidt (2003) and the references therein.

surprise higher wage offer and the possibility for the worker to reciprocate by exerting more effort—the information that the task matters to the researchers and that the “managers” stand to benefit from a job that is well done by receiving a bonus. Hence, we believe, the fact that gift-exchange is supposed to be at work is very salient for the workers. This, in turn, offers a good reason for why the initial wage gift is given and what the appropriate reaction to the gift is. In some of the above studies this is clear to employees due to the circumstances, but sometimes it is not [as in Gneezy and List (2006) where the wage increase comes as a surprise]; hence it is not clear why subjects should be contextually aware of the fact that they are engaged in a gift-exchange relationship and that exerting more effort is in fact appropriate.

Bellemare and Shearer (2009) analyze gift-exchange within a real firm (where the value of output is clear to the workers). In their study, there is a surprise bonus for the workers in a tree planting firm in British Columbia. Their results indicate a 10% increase in worker productivity on average which slowly dwindles. Moreover the effect of the gift is more marked if the worker has been with the firm for longer. Hence, Bellemare and Shearer argue that spot-market field experiments only establish a lower bound of the effects of gift-exchange in real firms that are characterized by longstanding and ongoing relations that amplify the effects. Based on data from the same firm, Bellemare and Shearer (2011) develop a structural behavioral model to identify a worker’s optimal response to monetary wage gifts. They use data from two separate field experiments to estimate the model and simulate how workers would react to different wage gifts. They find that profit-maximizing gifts would increase profits under slack labor market conditions by up to 10% on average.

The key innovation of the current study was to take seriously an immediate implication of reciprocity based gift-exchange models; the importance of the ability of the agent to repay the gift to the principal. According to the recent comprehensive study by DellaVigna et al. (2016), there are only three other studies, next that address this issue: Englmaier and Leider (2012b), Kessler (2013), and DellaVigna et al. (2016). Englmaier and Leider (2012b) also analyze the importance of the ability of the worker to “repay the gift” to the manager in a real-effort laboratory experiment where they vary the wage and the effect of the worker’s effort on the manager’s payoff. They report results that are consistent with the core findings of the current study plus additional predictions about which is the marginal worker (in terms of ability) affected by their experimental variation and how different types of individuals—selfish and reciprocal—react to it. Kessler (2013) uses a laboratory experiment and finds, consistent with our results, that gift-exchange is more prevalent when worker effort is more efficient. Finally, as part of a rich set of findings, DellaVigna et al. (2016) document in a large scale natural field experiment that the return of effort to the employer—as is predicted by this study or Englmaier and Leider (2012a)—positively affects the efficacy of gift-exchange.

The rest of the paper is organized as follows: The next section describes the design of the field experiment. Section 3 derives the theoretical predictions and Sections 4 and 5 present and discuss the results. Section 6 concludes. Appendices 1–3 contain illustrations, derivations and additional tables.

## 2 Experimental Design

The field experiment took place on the premises of the CLER lab at the Harvard Business School, where we provided subjects with computer work stations. We ran four sessions in September 2007. Though subjects were situated within one room, the layout was such that they could not monitor each other's work progress. Moreover, though we did not formally forbid communication, we did not note any signs of more than casual communication. Hence we conclude that there is only very limited scope for peer effects to affect our results. In total we had 59 participants; each of them worked for approximately five hours on a single day. The participants were hired via a temp worker agency that regularly works with the Harvard Business School.<sup>8</sup> We told the temp workers that we had been hired to organize a data entry project. These workers frequently work on similar data entry projects; hence there is no reason to believe that they suspected they were participating in a field experiment. Their job was to enter data from the 1849 Prussian Census into an Excel template.<sup>9</sup>

We created four treatment cells—*Low Wage/No Bonus*, *High Wage/No Bonus*, *Low Wage/Bonus*, *High Wage/Bonus*—based on the wage level and the bonus information. Workers were either paid the standard wage (\$13/h) as advertised by the temp agency, or were “surprised” with a higher wage (\$18/h).<sup>10</sup> In the baseline *No Bonus* treatments, we told the participants only that we were hired “by two professors to organize the entry of these data”. In the *Bonus* treatment groups we additionally inform the temps that we get a substantial “completion bonus” if enough work gets done.<sup>11</sup> These “bonus” treatments indicate whether the worker is willing to expend effort that benefits. Assignment of participants into one of the treatment cells was random and executed by the temp agency.

From the temp agency we get demographic information: the gender, race, age, work experience, and student status of the workers, which we use as controls. Most important, we have a measure of the workers' typing speeds (*Typing Score*), which we use to control for the temp workers' differing typing abilities. This is a key

<sup>8</sup> Atrium Staffing of Boston, <http://www.atriumstaff.com/>, described (as of 2007) their services as follows: “Since 1995, Atrium’s Office Support Practice has been the firm’s stronghold. Our team of skilled recruiters, who themselves draw from diverse career backgrounds, are experts in identifying administrative talent and understanding client needs. Our thorough screening processes ensure that we know our Associates well. Our customized approach to client searches results in impressive hire retention rates of over 93%. We staff direct hire, temp-to-hire and temporary administrative roles in a variety of companies.”

<sup>9</sup> An example from the 1849 Prussian Census can be found in Fig. 2 in Appendix 1.

<sup>10</sup> Note that in principle this is a setting where effort could be motivated by paying a piece rate as suggested by the classic moral hazard literature; cf. Holmström (1979). However, this is not done by the temp agency we work with. Hence, also below in the model, we will treat effort as not contractible (due to constraints on the contracting space).

<sup>11</sup> In order to not deceive the subjects on this issue we arranged with one of the providers of the Prussian Census Data, Sascha Becker, to put such an incentive scheme in place. Our actual “completion bonus” came in the form of free lunch and was worth roughly \$10.

determinant for productivity in this task. The final payment of the temp workers was done in cash directly at the end of each entry session.

Given that there are no previous papers that use this task or treatment variation, we were not able to calibrate the magnitude of the pay rise and the completion bonus in order to find the optimal combination to maximize gift-exchange; consequently our results should not be seen as an upper bound on the efficacy of reciprocal incentives. However, we are fairly certain that all participants considered the pay rise and completion bonus as “substantial”.

### 3 A Model of Reciprocal Motivation

Our experiment is not designed to differentiate exactly between different models of social preferences. Hence we do not interpret our findings as a strict test of our model, but rather consider the model to be a valuable frame within which to organize the data. We consider a simplified version of the model in Englmaier and Leider (2012a) where we solve the full moral hazard problem and derive the structure of the optimal contract in a standard principal agent problem with reciprocal agents. To lay out our model: We assume that there is a risk neutral manager who wants to maximize expected profits and one risk averse worker who cares about reciprocity. The worker can take an action (effort)  $a \geq 0$  with corresponding costs of effort  $c(a)$ , with:  $c'(a) > 0, c'(0) = 0$ ; and  $c''(a) > 0$ .

The actions imply a respective expected return for the manager  $ER(a)$  with  $ER'(a) > 0, ER''(a) \leq 0$ . In order to capture our experimental variation we introduce the scalar  $M$  which reflects the monetary value of output:  $M \cdot ER(a)$  is the expected monetary gross return for the manager from action  $a$ .

A contract  $(w, \hat{a})$  is a fixed wage payment  $w$ , as well as an unenforceable request for an action  $\hat{a}$ . In a real-world context we could think of  $\hat{a}$  as an informal job description or a code of conduct. In the experiment we will interpret  $\hat{a}$  as an exogenously given and commonly understood norm. Given our focus here on *changes in behavior* these details are not key to our results. While  $\hat{a}$  is not binding, it serves to fix the worker’s beliefs about the manager’s intended generosity (since the expected utility of a contract depends on the worker’s action).

The worker’s inherent concern for reciprocity is measured by  $\eta \in [0, +\infty)$ . The worker’s utility function—given that she takes action  $a$ , under the contract  $(\tilde{w}, \hat{a})$ —is given by

$$U(\tilde{w}, a, \hat{a}) = u(\tilde{w}) - c(a) + \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u}) \cdot M \cdot (ER(a) - \tilde{w}), \tag{1}$$

where  $\bar{u}$  is the worker’s outside option in the labor market. The utility function captures—albeit in a simplistic form—the core idea of reciprocal motivation: If an individual has been treated kindly, she will want to reciprocate in kind.<sup>12</sup> The function

<sup>12</sup> Rabin (1993), for simultaneous move games, and Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006), for sequential games, have developed powerful and general models of reciprocal preferences.

consists of three parts: (i) utility from the monetary wage payment  $u(\tilde{w})$ ; (ii) effort costs  $c(a)$ ; and (iii) reciprocal utility  $\eta(u(\tilde{w}) - c(\hat{a}) - \bar{u})M \cdot (ER(a) - \tilde{w})$ , where  $\eta$  measures the intensity of the reciprocal preferences.

A “generous” contract is one that provides a rent to the worker: an expected monetary utility in excess of the worker’s outside option. A more generous contract will induce the worker to feel more reciprocal, which here means that she will derive greater marginal and absolute utility from the manager’s profit. On the assumption that the contract is generous, the worker’s optimal effort choice  $a^*$  for a given contract is implicitly defined by the first order condition

$$\frac{\partial U(\tilde{w}, a, \hat{a})}{\partial a} = -c'(a^*) + \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u})M \cdot ER'(a^*) = 0. \tag{2}$$

Applying the implicit function theorem we can derive the relevant comparative statics w.r.t.  $\tilde{w}$  and  $M$ : They are positive, as is the cross partial w.r.t.  $\tilde{w}$  and  $M$ , which indicates that they are complements.<sup>13</sup> Note that—since  $M$  and  $\eta$  (the concern for reciprocity) always appear together—the effect of varying  $\eta$  is the same as the effect of varying  $M$ . The following Lemma 1 summarizes these results.

**Lemma 1 (Reciprocity)** *For a generous contract,  $(u(\tilde{w}) - c(\hat{a}) - \bar{u}) > 0$ , the worker’s optimal action  $a^*$  is implicitly defined by the first order condition*

$$\frac{\partial U(\tilde{w}, a, \hat{a})}{\partial a} = -c'(a^*) + \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u}) \cdot M \cdot ER'(a^*) = 0. \tag{3}$$

*It is increasing in  $\tilde{w}$ :  $\partial a^* / \partial \tilde{w} > 0$ ; increasing in  $M$ :  $\partial a^* / \partial M > 0$ ; increasing in  $\eta$ :  $\partial a^* / \partial \eta > 0$ ; and  $\tilde{w}$  and  $M$  are complements:  $\partial^2 a^* / \partial \tilde{w} \partial M > 0$ .*

The intuition for the complementarity is fairly straightforward from the utility function: Increasing the wage leaves a larger rent to the worker and increases the weight that she gives to the managers welfare. Due to the multiplicative structure, the worker finds it more attractive to work harder when she has a stronger impact on the manager’s surplus.

In contrast to the standard model of preferences ( $\eta = 0$ ) the first-order condition simplifies to

$$\frac{\partial U(\tilde{w}, a, \hat{a})}{\partial a} = -c'(a^*) < 0, \tag{4}$$

which is always negative, since with flat wages the worker’s utility unambiguously decreases in her effort choice and her optimal action  $a^*_{standard}$  is trivially given by  $a^*_{standard} = 0$ . Increasing  $M$  or  $\tilde{w}$  has no effect on  $a^*_{standard} = 0$ . This is summarized in Lemma 2:

<sup>13</sup> See Appendix 2 for the derivations and a parameterized example.



**Lemma 2** (Standard Preferences).  $a^*_{Standard} = 0$ , and the corresponding comparative statics are trivially given by

$$\frac{\partial a^*_{standard}}{\partial M} = 0, \quad \frac{\partial a^*_{standard}}{\partial \tilde{w}} = 0, \quad \frac{\partial^2 a^*_{standard}}{\partial M \partial \tilde{w}} = 0.$$

Summarizing the results from our model, we can formulate the following predictions that we try to validate in our experimental analysis. The first three follow directly from Lemma 1:

**Prediction 1** *Effort is increasing in wage  $\tilde{w}$ .*

**Prediction 2** *Effort is increasing in managerial payoff  $M$ .*

**Prediction 3** *Managerial payoff and wage are complementary in their effect on effort,  $\partial a^{*2} / \partial \tilde{w} \partial M > 0$ .*

The next prediction is straightforward: reciprocal incentives via gift-exchange work better for more reciprocally inclined subjects. In particular, low reciprocity individuals should be unlikely to exhibit a positive response to high wages, as their utility from providing effort for the principal is very low compared to their cost of effort. High-reciprocity individuals, however, are more likely to have a sufficiently strong utility benefit from returning the gift so as to induce effort.

**Prediction 4** *The positive effort response to a wage gift or a manager bonus is more pronounced for more reciprocally inclined (higher  $\eta$ ) subjects.*

## 4 Experimental Results

Our main performance measure are the subjects' data entry rates: the number of characters of data entered per minute.<sup>14</sup> As a robustness check, in our regressions we also report the accuracy-corrected data entry rate.<sup>15</sup> In Table 1 we present the mean and median entry rates for the whole shift. Subjects' overall productivity suggests that offering a high wage in the *No Bonus* treatments had a negative impact on

<sup>14</sup> The Prussian Census has two volumes that differ slightly in their content: in particular, in the number of cells/line and the average number of characters/cell. To control for these differences we convert the number of cells (which is our basic measure of performance) into characters entered by weighting cells by the average number of characters/cell in the respective volume. Alternatively, we can also directly control for the volume by adding a dummy variable and keeping cells as the measure. These results are quantitatively and qualitatively similar.

<sup>15</sup> To construct this measure we repeat the exercise used to construct the data entry rate, i.e. we convert the number of correctly entered cells into characters entered by weighing cells with the average number of characters/cell in the respective volume.

**Table 1** Data-entry rate (Chars/min) by treatment

Employee wage	Manager bonus	Data entry rate			Accuracy corrected rate			# of Obs.
		Mean	Median	SD	Mean	Median	SD	
\$13/h	No	29.71	25.52	16.12	28.02	23.65	15.32	14
\$18/h	No	23.47	19.83	11.47	21.92	19.42	10.57	15
\$13/h	Yes	24.38	23.40	9.21	23.06	22.09	9.11	15
\$18/h	Yes	28.36	29.25	8.42	26.72	27.48	8.24	15

The table shows the average data-entry rate, measured as characters/minute (Data-Entry Rate) or as correct characters/minute (Accuracy-Corrected Rate), as well as the median rate and the standard deviation classified by treatment

**Table 2** Typing score

Employee wage	Manager bonus	Typing score			# of Obs.
		Mean	Median	SD	
\$13/h	No	57.93	52.50	17.46	14
\$18/h	No	48.80	46.00	13.97	15
\$13/h	Yes	48.27	48.00	14.35	15
\$18/h	Yes	56.20	53.00	16.83	15

The table shows the average Typing Score as provided by the temp agency, as well as the median rate and the standard deviation broken up by treatment

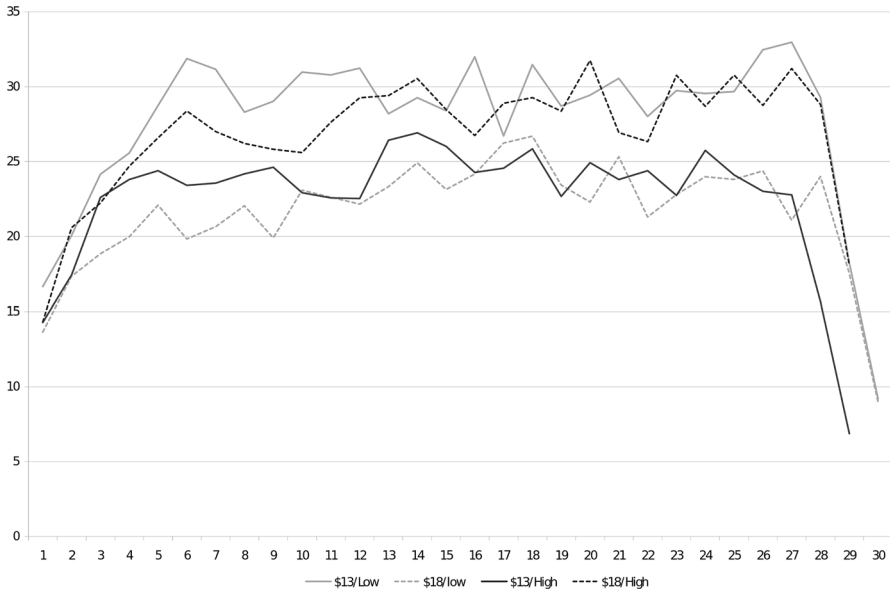
effort, while it had a positive effect in the *Bonus* treatments. We test this effect statistically in the analysis that follows.

Figure 1 shows how productivity—measured in 10 minute intervals—evolves over time in the four treatment conditions. Though the levels differ, there seems to be no difference in the time trend. There is learning in all four treatments for the first hour, then productivity is fairly flat until the end, where there is a steep drop in all treatments. This drop is mostly 'technical': Subjects finished a line and did not start a new one that they were unlikely to finish before the end of their shift.

As can be seen in Table 2 there are considerable differences in underlying ability (the typing speed score) between treatments. Hence, a direct comparison of the overall productivity between treatments is somewhat misleading. We address this below by directly controlling for ability in our main regressions.

#### 4.1 Regression Analysis of Treatment Effects

Our sample size in terms of participating subjects is only somewhat bigger than in preceding studies: e.g., Gneezy and List (2006) or Bellemare and Shearer (2009); but importantly, our design allows us to use detailed additional information to estimate more precisely the treatment effects. Most importantly we have a measure of data-entry ability from the temp agency. The agency has temp workers take a typing speed test



**Fig. 1** Performance over Time. The x-axis shows the 10 minute sub-periods and the y-axis shows the average data entry rates (characters/minute) by treatment. \$13/Low refers to the \$13 Wage/No Bonus treatment., \$18/Low refers to the \$18 Wage/No Bonus treatment, etc.

upon hiring, and the agency made this information available to us. Since Table 2 indicates substantial heterogeneity in the underlying ability distribution across treatments we proceed by including controls for worker ability. Additionally, we have worker productivity in our task at 10-min and 30-min intervals. We will use the 10-min data as the units of observation in this paper. The results for the 30-min data are quantitatively similar but less precisely estimated and are available upon request.<sup>16</sup> Table 3 presents the results of a GLS estimate with a heteroskedastic panel structure and AR(1) errors. A Wooldridge test for serial correlation finds significant autocorrelation ( $p < 0.001$ ) while a Likelihood Ratio test suggests panel heteroskedasticity ( $p < 0.001$ ).

The two alternate effort measures presented in specifications (1) and (2) in Table 3 give qualitatively and quantitatively very similar estimates, which suggests that the response to the treatment variations does not affect quality but only influences the quantity margin. Contrary to the fundamental gift-exchange intuition, we find a significant and negative effect of the wage gift on effort and no significant effect of the manager bonus. However, the interaction *High Wage* × *Manager Bonus* is significantly positive and large in size, which results in an overall positive

<sup>16</sup> Gneezy and List (2006) and Kim and Slonim (2009) use the same units of observation. Our performance measure is constructed from information when whole lines of the census were finished. These lines took several minutes to be completely entered. To smooth our data, we proceed as follows: If a line was started  $x$  minutes before the end of the  $n$ -th 10 min sub-period and finished in the  $y$ -th minute of the  $(n+1)$ -th 10 min sub-period, we account  $x/(x+y)$  of the characters of this line to the  $n$ -th and  $y/(x+y)$  of the characters of this line to the  $(n+1)$ -th 10 min sub-period.

**Table 3** Regression: performance (10-min. Periods)

Variables	(1)	(2)
	Data entry rate	Accuracy corrected rate
Typing speed test score	0.220*** (0.0272)	0.203*** (0.0259)
Temps paid high wage	- 3.531*** (1.118)	- 3.355*** (1.063)
Manager bonus	- 1.526 (1.357)	- 0.994 (1.303)
High wage × manager bonus	6.456*** (1.731)	5.824*** (1.665)
Period	0.105*** (0.0359)	0.0700** (0.0344)
Demographic controls	Yes	Yes
Constant	17.46*** (5.266)	11.71** (5.006)
Observations	1628	1628
Number of subjects	57	57
Total effect: high wage and manager bonus	2.925** (1.214)	2.469** (1.168)

We report cross-sectional time-series FGLS regression with heteroskedastic panel structure and AR(1) correlation. Demographic Controls include dummy variables for gender, race, age, work experience, and student status. Standard errors are reported in parentheses. Significance is denoted: \*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

gift-exchange in the *Bonus* treatment: The effect of a wage gift depends importantly on the characteristics of the job context, and there is a strong complementarity between the wage gift and the magnitude of the managerial payoff.<sup>17</sup>

## 4.2 Treatment Effects by Agreeableness

As we have information on workers' scores in a "Big 5" personality test, we can analyze how effects vary along this dimension. The "Big 5" personality factors capture five distinct aspects of an individual's personality: *extraversion*; *agreeableness*; *conscientiousness*; *emotional stability* (also referred to as its inverse, *neuroticism*); and *imagination* (or *openness to experience*). The five factors are typically measured by having subjects read a series of statements, such as "I pay attention to details", and rate them on a scale from "strongly agree" to "strongly disagree", where each statement is associated with one of the factors. Extraversion represents an individual's sociability and engagement

<sup>17</sup> Relaxing our specification by allowing for non-linear effects of ability, we estimate a specification with separate dummy variables for the lowest, middle, and highest terciles of ability. In this analysis, we find very similar coefficients and present them in Table 7 in Appendix 3. Note that the results in Table 7 suggest non-linear effects of ability with the middle tercile not being different from the top tercile but the bottom tercile being significantly slower.

**Table 4** Summary the personality traits of the subjects

Personality factor	Concepts	Mean	SD	Min	Max
Extraversion	Sociability, seeks fulfillment from others	35.87	7.48	18	50
Agreeableness	Adapting to suit others, politeness	41.64	4.48	32	50
Conscientiousness	Honesty, rule following, diligence	34.59	6.46	19	50
Emotional	Stability calm, not easily upset	31.44	7.13	16	47
Imagination	Seeks new experiences and intellectual pursuits	38.39	5.22	27	50

Participant summary statistics based on a standard 50 question Big 5 personality test administered separately after the primary data entry task

with the external world. Extraverts are commonly seen as energetic and action oriented. Extraversion is positively associated with scale items such as “I am the life of the party” and negatively associated with items such as “I don’t talk a lot.” Agreeableness represents an individual’s interest in social harmony and getting along with others. Agreeable individuals are seen as helpful, trusting and optimistic. Agreeableness is positively associated with items such as “I take time out for others” and negatively associated with items such as “I am not interested in other people’s problems.” Conscientiousness represents an individual’s self-discipline and inclination to follow rules and instructions. Conscientious people are seen as honest, hardworking, able to control their impulses, and inclined to plan. Conscientiousness is positively associated with scale items such as “I am always prepared” and negatively associated with items such as “I shirk my duties.” Emotional stability reflects an individual’s calmness and low propensity to feel negative emotions. Conversely, neurotic individuals are emotionally reactive to stress and frustration, inclined towards a bad mood and anxiety. Emotional stability is positively associated with scale items such as “I am relaxed most of the time” and negatively associated with items such as “I worry about things.” Finally, imagination represents an individual’s attraction to unusual ideas, adventure, artistic expression and a variety of experiences. Imaginative people are seen as intellectually curious, sensitive to beauty, and willing to try new things. Imagination is positively associated with scale items such as “I am full of ideas” and is negatively associated with scale items such as “I am not interested in abstractions.” We use a standard 50 question scale, and Table 4 presents summary statistics for our sample. Our results are within the typical range and are similar to the results for a standard laboratory subject pool in our companion paper; Englmaier and Leider (2012b). For comparability between factors, in all of the analyses below we use z-scores to standardize each factor.

In our study we identify subjects who score highly on the trait “agreeableness”, which had been shown experimentally to relate to standard lab measures of reciprocity; cf. Ben-Ner et al. (2004) or Ashton et al. (1998). While Ben-Ner et al. (2004) and Ashton et al. (1998) also find some evidence that “openness” and “emotional stability” may relate to reciprocity as well, the relationship between reciprocity and agreeableness was robust across specification and sample. The results in Englmaier and Leider (2012b) also show this correlation. Empirical evidence suggests that agreeable individuals not only reciprocate more, but are also more trustworthy and altruistic. However, our model of reciprocal workers is to be understood as an illustration of one particularly relevant motivational mechanism in our context.

As opposed to the trust game or other lab measures of reciprocity, personality tests such as the “Big 5” are quite common in the hiring practices of firms. In particular, high agreeableness corresponds with one of the criteria that Autor and Scarborough (2008) identify in the hiring practice of the firm that they study, where the firm gave hiring preference to applicants with positive z-scores for agreeableness, conscientiousness, and extroversion. The separation of our subjects along the agreeableness dimension—which proxies for a split by reciprocity—allows us to examine whether our effects are in fact driven by the reciprocal subjects, as is suggested by our model. Exploiting the population heterogeneity along the reciprocity dimension and in doing so testing a key tenet of the reciprocity models was generally not done in early studies on gift-exchange. This was generally due to a lack of data on the preference. Examples of more recent work that pays attention to this heterogeneity and that finds support for heterogeneous treatments effects in line with reciprocity models are Englmaier and Leider (2012b) and Cohn et al. (2015).

When we turn to the results of our study, first note that there is no correlation between our agreeableness measure and the Typing Speed Test Score (Spearman  $\rho = 0.0482$ ,  $p = 0.7170$ ). Table 5 shows our basic regression analysis for the sample split into highly agreeable (strongly reciprocal; top third) and low agreeable (weakly reciprocal; bottom two thirds) workers.<sup>18</sup> Again, there are substantial differences in the response to the treatment variation across the agreeableness dimension. The negative response to the wage gift in the *No Bonus* treatment is driven entirely by the low agreeableness workers. Additionally, the low agreeableness workers have no overall response to the wage gift in the *Bonus* treatment. However, note that the complementarity effect between higher wage and manager bonus is significantly positive and substantial for the low agreeableness group. While at first glance puzzling, this result is in line with theoretical arguments and empirical findings in Englmaier and Leider (2012b) and Englmaier et al. (2014): intuitively, low-agreeableness workers are not (necessarily) unresponsive to gift exchange; however, they need a bigger boost to react—as is provided by giving both the wage gift and the improved ability to reciprocate by the manager bonus. In contrast, high-agreeableness workers react already to either one in separation as indicated by the (admittedly noisily estimated) coefficients in the regressions. These high-agreeableness workers have also a significantly positive response in the *Bonus* treatment (indicated by the significant total effect reported at the bottom of Table 5). Therefore, high-agreeableness workers can be induced to exert significantly more effort than the control low-wage treatment; but they have in some sense a more gradual reaction to the changes in the work setting as compared to low agreeableness workers who react primarily to the combination of high wage and manager bonus. Overall, our results are consistent with our model that a response to a wage gift should come primarily from reciprocal workers. What remains puzzling is the strong negative effect of the wage gift among the low-agreeableness subjects. We will return to this topic in Sect. 5.

<sup>18</sup> The qualitative results are robust to using different cut offs for classifying subjects as highly agreeable. Directly interacting the continuous agreeableness score with the interactions of interests is somewhat too demanding given our sample size.

**Table 5** Regression: performance (10-min. Periods) by agreeableness score

Variables	Data entry rate		Accuracy corrected rate	
	Low Agr. (1)	High Agr. (2)	Low Agr. (3)	High Agr. (4)
Typing-speed test score	0.226*** (0.0502)	0.0590 (0.0531)	0.188*** (0.0490)	0.0770 (0.0500)
Temps paid high wage (\$18/h)	- 4.142** (1.681)	2.392 (2.675)	- 4.294*** (1.614)	2.159 (2.535)
Manager bonus	- 0.377 (1.799)	4.856 (3.004)	- 0.167 (1.745)	4.224 (2.870)
High wage × manager bonus	5.380** (2.495)	2.041 (2.676)	5.222** (2.408)	1.332 (2.562)
Period	0.0819* (0.0433)	0.134** (0.0553)	0.0522 (0.0419)	0.0898* (0.0527)
Demographic controls	Yes	Yes	Yes	Yes
Constant	16.94** (7.113)	67.05*** (14.76)	12.65* (6.775)	58.30*** (13.93)
Observations	1059	569	1059	569
Number of subjects	37	20	37	20
Total effect: high wage and manager bonus	1.237 (1.598)	4.433** (2.165)	0.927 (1.560)	3.491* (2.058)

“Agr.” abbreviates “Agreeableness”. Highly Agreeable subjects are those with an agreeableness score in the top 33%. We report cross-sectional time-series FGLS regression with heteroskedastic panel structure and AR(1) correlation. Demographic controls include dummy variables for gender, race, age, work experience, and student status. Standard errors are reported in parentheses. Significance is denoted: \*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

### 4.3 Treatment Effects over Time

Finally, we consider the question of how productivity evolves over time. As Gneezy and List (2006) document a fast disappearing gift-exchange pattern, we are in particular interested in how the effects of the treatment variations evolve over the course of the experiment. We split our data into first- and second-half observations and report in Table 6 the base regressions separately. Again, the qualitative results are robust to different divisions of the data.<sup>19</sup> For both treatments the effect of a wage gift has in fact a *more positive* effect in the second half of the experiment. The negative response to the wage in the *No Bonus* treatment only exists in the first half; it disappears in the second half. Similarly, the positive effect in the *Bonus* treatment is present only in the second half of the experiment. However, the strong complementarity between the treatments—*High Wage × Manager Bonus*—is very stable over time. This suggests that the positive response to gift-exchange does not have to be a short-lived phenomenon.

<sup>19</sup> E.g., performing an hourly based analysis gives similar results.

**Table 6** Regression: performance (10-min. Periods) by 1st/2nd half

Variables	Data entry rate		Accuracy corrected rate	
	1st half (1)	2nd half (2)	1st half (3)	2nd half (4)
Typing-speed test score	0.216*** (0.0320)	0.258*** (0.0360)	0.200*** (0.0299)	0.241*** (0.0341)
Temps paid high wage (\$18/h)	- 4.795*** (1.226)	- 1.527 (1.521)	- 4.319*** (1.183)	- 1.581 (1.423)
Manager bonus	- 0.572 (1.407)	- 1.249 (1.856)	- 0.429 (1.371)	- 0.460 (1.778)
High wage × manager bonus	5.249*** (1.902)	5.406** (2.334)	4.823*** (1.823)	4.531** (2.237)
Period	0.578*** (0.0656)	- 0.372*** (0.0984)	0.532*** (0.0630)	- 0.430*** (0.0930)
Demographic controls	Yes	Yes	Yes	Yes
Constant	15.91*** (5.914)	25.27*** (7.458)	10.67* (5.673)	17.19** (7.022)
Observations	855	773	855	773
Number of subjects	57	57	57	57
Total effect: high wage and manager bonus	0.453 (1.324)	3.878** (1.623)	0.504 (1.287)	2.949* (1.537)

We report cross-sectional time-series FGLS regression with heteroskedastic panel structure and AR(1) correlation. Demographic Controls include dummy variables for gender, race, age, work experience, and student status. Standard errors are reported in parentheses. Significance is denoted: \*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

## 5 Discussion of the Results

Previous experiments on gift-exchange in the field studied a variety of job situations (e.g., one-time jobs versus ongoing jobs); job tasks (e.g., data entry, fund raising, and tree planting); and worker pools (e.g., students versus full-time employees). Our results suggest that it may not be surprising that the observed effect of surprise wage increases varies across these studies.<sup>20</sup>

We demonstrate that a positive gift-exchange can exist, but that its presence depends on the characteristics of the job. Therefore we should not expect gift-exchange to be present in all job settings. In particular, the results from our field study suggest that the extent to which the manager will directly benefit when subordinates produce high output plays an important role for the efficacy of reciprocity and gift-exchange in the field. We therefore anticipate that gift-exchange should be most prominent in settings where managers have a strong direct benefit from employee effort, e.g. if they have performance-based monetary incentives.

<sup>20</sup> Kessler (2013) and Esteves-Sorenson (2018) are two careful and comprehensive papers that explicitly try to explain the heterogeneous treatment effects in the various extant studies. Moreover, they examine other characteristics of the work setting that determine the effectiveness of reciprocity-based incentives.



An analysis of the UK *Workplace Employment Relations Study* (WERS) by Englmaier et al. (2016) finds some support for this. Now with increasingly better available representative administrative data on HR practices it would be interesting to further study the richer predictions of reciprocity-based models for organizational design; cf. Englmaier and Leider (2012a).

The negative effect on effort in our *No Bonus* treatment is puzzling. It is possible that absent an indication that high effort is important to the manager the surprise high wage appears wasteful. Subjects might interpret the out-of-context wage increase as a mistake on our side, which would cast doubt on our managerial aptitude. From this they might have downwards updated also our ability or willingness to “monitor their work”. This indicates that the manager who has incentives that are related to output may put the decision to offer a higher wage into context and avoid this negative updating on aptitude.

While we did not *ex ante* predict that the effect of a gift would get stronger over time, a few explanations are possible: First, we observe some level of learning in all treatments. It may take some time before workers learn their cost of effort, which determines how willing they are to exert effort to provide a return gift to the principal. Additionally, perhaps gift-exchange raises productivity by increasing persistence, rather than maximal effort. As previously mentioned, we see no treatment effect on quality, which suggests that workers need not improve on all dimensions of effort. Last, workers may have been waiting to see if another surprise was coming before completely determining their effort level.

As in previous experiments—see for example DellaVigna et al. (2016)—we do not find that the increase in worker wages “pays for itself” in increased productivity: the 40% increase in wage garnered only a 10% increase in productivity. However, we were using completely flat wages to pay the workers. The theoretical model in Englmaier and Leider (2012a) suggests that the level of reciprocity that is needed to induce effort under a flat wage is generally larger than the amount of reciprocity that can reduce agency costs when paired with explicit monetary incentives. A flat wage is generically far from the optimal contract, and so it is quite possible that any given level of reciprocity could be much more powerfully leveraged with a more sophisticated incentive scheme. Bellemare and Shearer (2009) find that a fixed gift of \$80 (40% of earnings) on top of the standard piece rate increased productivity by an amount worth \$40. As in our study the size of the gift was chosen arbitrarily; hence fine tuning the mix of gifts and explicit incentives is likely to be critical. Going forward, establishing that gift-exchange works not only conceptually, but also can be profit-enhancing is one of the key challenges for this literature.

## 6 Conclusion

The importance of fairness and social preferences—especially for the work relationship—has long been documented in the lab. However, a number of studies have highlighted that the empirical importance of reciprocity as a motivating force in the field crucially depends on details of the environment. The wage gift

has to stand in the proper context to generate a motivating effect. Based on a principal-agent model of reciprocal motivation, we argue that a key determinant of the efficiency of reciprocal motivation is the ability of an agent to repay a gift: The benefit accrues to the principal from high effort.

Our results confirm the importance of the specific characteristics of the job-setting in generating gift-exchange-based incentives. If the manager stands to benefit much from additional effort, there is a strong positive effort response; while in the absence of the managerial benefit, we find no or even a negative response to the wage gift. This indicates a strong complementarity between the wage gift and the managerial payoff in generating incentives for employee effort.

Additionally, we find that the positive response to high wages comes primarily from high-reciprocity workers (as identified by a personality test), while the negative response comes from low-reciprocity workers. We find no weakening of positive responses to wage gifts over time. To the contrary, any negative effects of the wage gift disappear in the later stages of the task and we find an overall strongly positive effect of our treatment manipulations.

Our study indicates that employing agents' reciprocity as a part of a firm's personnel policy is a viable alternative and can be successfully done. However, as highlighted by, e.g., Ichniowski and Shaw (2003) or Bartling et al. (2012), it is important that various complementary parts of the firm's compensation and HR policy are coordinated to maximize the effect of reciprocity. For example, firms that wish to employ reciprocal incentives may want to select for reciprocal motivations during hiring. An interesting implication of our study is that performance-related pay for middle management should be part of a remuneration policy even absent moral-hazard problems at that hierarchy level.

Our results suggest several avenues for future research: Further empirical work could identify the optimal magnitude of the gift and the proper mix between reciprocal and explicit motivation to maximize the profitability of gift-exchange.

**Acknowledgements** We thank Greg Baron, Gary Charness, Tore Ellingsen, Lena Greska, Al Roth, Carmit Segal, Caspar Siegert, Christina Strassmair, Victor Tremblay, Mo Xiao, and participants of the 2011 EWEBE meeting for their comments and suggestions. We also thank Sascha Becker and Ludger Wössmann for providing us with the Prussian Census materials for our field experiment. Florian Englmaier thanks the Harvard Business School for its hospitality and the German Science Foundation (DFG) for financial support via grants *EN 784/2-1*, *SFB/TR-15*, and *SFB/TR-190*. Steve Leider thanks the National Science Foundation and the Sperry Fund for financial support. The study received IRB approval by Harvard University's Committee on the Use of Human Subjects in Research, Application Numer: F15221-101.

## Appendix 1: Additional Material

See Fig. 2.

Namen der Kreise.	Städte und plattes Land.	Gebäude.								
		Oeffentliche Gebäude.						Privat-Gebäude.		
		Zum öfentlichen Gebrauche bestimmte Küchen und Bekäuser.	Schulkäuser für den öfentlichen Unterricht.	Zur Aufnahme und Verpflegung von Wägen, Kutschen, Atzen, schwarzen und weißbaren Personen bestimmte Gebäude.	Gebäude zur Verammlung und Kreis- führung der Landeskolle- gien, Justiz, Polizei, Steuers-Bü- reau, Magistrate und Gemein- de-Vorstände.	Zu andern Zwecken der öfentlichen, Civil- und Comunal- Behörden und Anstalten bestimmte Gebäude.	Militär- Gebäude mit Einzelnen der verschieden alten Art bestimmte, sowie auch die Militär- Lazarethe.	Privat- Wohn- häuser.	Fabrik- Gebäude, Möhlen und Privat- Magazine.	Ställe, Scheunen und Schoppen.
<b>I. Regierungsbezirk Königsberg.</b>		1.	2.	3.	4.	5.	6.	7.	8.	9.
Memel.	Memel .....	5	9	6	6	15	4	792	201	355
	Plattes Land ..	5	47	2	7	5	"	3,494	116	6,259
	Summe	10	56	8	13	20	4	4,286	317	6,614
Fischhausen.	Pillau .....	1	3	"	4	15	46	183	17	36
	Fischhausen ..	1	1	1	1	4	"	140	17	216
	Plattes Land ..	14	64	1	4	30	"	2,757	111	3,230
	Summe	16	68	2	9	49	46	3,080	145	3,482
Königsberg Stadt.	Königsberg ...	20	33	65	18	112	76	4,044	707	2,281
Königsberg Land.	Plattes Land ..	18	77	7	3	18	"	3,043	189	3,718
Labiau.	Labiau .....	1	2	1	1	2	"	240	30	412
	Plattes Land ..	6	66	"	7	4	"	4,472	133	3,883
	Summe	7	68	1	8	6	"	4,712	163	4,295
Welan.	Welan .....	1	2	3	1	9	4	250	18	344
	Tapiau .....	2	1	6	"	1	"	149	3	160
	Allenberg ....	1	1	2	1	5	"	190	12	180
	Plattes Land ..	8	62	1	2	3	1	3,051	195	3,790
	Summe	12	66	12	4	18	5	3,640	228	4,474
Gerdaun.	Gerdaun .....	2	4	2	"	7	"	204	11	251
	Nordenburg ..	1	1	2	1	6	"	193	11	296
	Plattes Land ..	10	61	4	"	11	"	2,249	116	3,156
	Summe	13	66	8	1	24	"	2,646	138	3,703
Rastenburg.	Drengfurt ....	1	2	1	1	"	"	168	3	262
	Rastenburg ...	3	4	1	1	9	3	355	5	254
	Barten .....	1	1	2	1	4	"	127	30	225
	Plattes Land ..	13	79	12	1	4	"	2,372	121	3,328
	Summe	18	86	16	4	17	3	3,022	159	4,069
Friedland.	Bartenstein ...	2	2	"	"	13	3	317	3	351
	Friedland .....	1	2	3	3	5	"	241	2	324
	Schuppenbeil ..	1	4	1	1	4	"	229	22	155
	Domnau .....	1	2	1	"	4	"	132	2	238
	Plattes Land ..	13	57	7	3	2	"	2,061	134	3,181
	Summe	18	67	12	7	28	3	2,980	163	4,249
Prouss. Eilau.	Pr. Eilau .....	2	3	2	4	1	"	196	2	395
	Landsberg ...	1	1	1	1	"	"	225	4	268
	Creutzburg ...	1	3	1	2	3	"	218	2	287
	Plattes Land ..	19	81	4	1	8	"	3,520	162	4,325
	Summe	23	88	8	8	12	"	4,159	170	5,265
Heiligenbeil.	Heiligenbeil ..	1	1	2	1	3	"	318	21	433
	Zinten .....	1	2	2	1	3	"	258	5	219
	Plattes Land ..	15	66	13	3	22	"	3,088	169	3,539
	Summe	17	69	22	5	28	"	3,664	193	4,191

Fig. 2 Example Page Prussian Census of 1849

## Appendix 2: Derivations

### Comparative Statics

On the assumption that the contract is generous, the worker’s optimal effort choice  $a^*$  for a given contract is implicitly defined by the first order condition

$$\frac{\partial U(\tilde{w}, a, \hat{a})}{\partial a} = -c'(a^*) + \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u})M \cdot ER'(a^*) = 0.$$

Applying the implicit function theorem we can derive the relevant comparative statics w.r.t.  $\tilde{w}$  and  $M$ :

$$\begin{aligned} \frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a^2} &= -c''(a^*) + \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u})M \cdot ER''(a^*) < 0 \\ \frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a \partial M} &= \eta(u(\tilde{w}) - c(\hat{a}) - \bar{u})ER'(a^*) > 0 \\ \frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a \partial \tilde{w}} &= \eta M \cdot u'(\tilde{w}) \cdot ER''(a^*) > 0 \end{aligned}$$

Hence it follows that

$$\frac{\partial a^*}{\partial M} = -\frac{\frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a \partial M}}{\frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a^2}} > 0; \quad \frac{\partial a^*}{\partial \tilde{w}} = -\frac{\frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a \partial \tilde{w}}}{\frac{\partial^2 U(\tilde{w}, a, \hat{a})}{\partial a^2}} > 0; \quad \text{and} \quad \frac{\partial^2 a^*}{\partial M \partial \tilde{w}} > 0.$$

### A Simple Example

To provide a concrete example assume that

$$ER(a) = a; \quad u(\tilde{w}) = \tilde{w}; \quad \text{and} \quad c(a) = \frac{a^2}{2}.$$

Then the worker’s utility can be written as

$$U(\tilde{w}, a, \hat{a}) = \tilde{w} - \frac{a^2}{2} + \eta \left( \tilde{w} - \frac{\hat{a}^2}{2} - \bar{u} \right) M(a - \tilde{w}),$$

the first order condition is given by

$$-a + \eta \left( \tilde{w} - \frac{\hat{a}^2}{2} - \bar{u} \right) M = 0,$$

and can be rearranged to define explicitly  $a^*$

$$a^* = \eta \left( \tilde{w} - \frac{\hat{a}^2}{2} - \bar{u} \right) M.$$

Now taking explicit derivatives gives

$$\frac{\partial a^*}{\partial M} = \eta \left( \tilde{w} - \frac{\hat{a}^2}{2} - \bar{u} \right) > 0; \quad \frac{\partial a^*}{\partial \tilde{w}} = \eta M > 0; \quad \text{and} \quad \frac{\partial^2 a^*}{\partial M \partial \tilde{w}} = \eta > 0.$$

### Appendix 3: Additional Analysis

See Table 7.

**Table 7** Regression: performance (10 min. Periods) incl.

Variables	(1) Data entry rate	(2) Accuracy corrected rate
Slowest third	- 5.451*** (0.987)	- 4.836*** (0.945)
Fastest third	- 0.407 (1.075)	- 0.191 (1.038)
Temps paid high wage	- 3.541*** (1.131)	- 3.324*** (1.074)
Manager bonus	- 2.506* (1.340)	- 1.875 (1.294)
High wage × manager bonus	7.853*** (1.700)	7.048*** (1.633)
Period	0.122*** (0.0369)	0.0867** (0.0354)
Demographic controls	Yes	Yes
Constant	15.94*** (5.488)	10.35** (5.236)
Observations	1628	1628
Number of subjects	57	57
Total effect: high wage and manager bonus	4.312*** (1.215)	3.724** (1.177)

We report cross-sectional time-series FGLS regression with heteroskedastic panel structure and panel-specific AR(1) correlation. Demographic controls include dummy variables for gender, race, age, work experience, and student status. Standard errors are reported in parentheses. Significance is denoted: \*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

### References

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

Ashton, M., Paunonen, S., Helmes, E., & Jackson, D. (1998). Kin altruism, reciprocal altruism, and the Big Five personality factors. *Evolution and Human Behavior*, 19, 243–255.

Autor, D., & Scarborough, D. (2008). Does job testing harm minority workers? Evidence from retail establishments. *Quarterly Journal of Economics*, 123(1), 219–277.

- Bartling, B., Fehr, E., & Schmidt, K. M. (2012). Screening, competition, and job design: Economic origins of good jobs. *American Economic Review*, *102*(2), 834–864.
- Becker, S., Messer, D., & Wolter, S. (2012). A gift is not always a gift: Gift exchange in a voucher experiment. *Economica*, *80*(318), 345–371.
- Bellemare, C., & Shearer, B. (2009). Gift giving and worker productivity: Evidence from a firm level experiment. *Games and Economic Behavior*, *67*, 233–244.
- Bellemare, C., & Shearer, B. (2011). On the relevance and composition of gifts within the firm: Evidence from field experiments. *International Economic Review*, *52*(3), 855–882.
- Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, *25*(5), 581–589.
- Cohn, A., Fehr, E., & Götte, L. (2015). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, *61*(8), 1777–1794.
- DellaVigna S., List J., Malmendier U., & Rao G. (2016). *Estimating social preferences and gift exchange at work*. NBER Working Paper No 22043 .
- Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, *47*, 268–298.
- Englmaier, F., & Leider, S. (2012a). Contractual and organizational structure with reciprocal agents. *American Economic Journal: Microeconomics*, *4*(2), 146–183.
- Englmaier F., & Leider S. (2012b). *Gift exchange in the lab—It is not (only) how much you give...* CESifo Workingpaper 2944.
- Englmaier, F., Strasser, S., & Winter, J. (2014). Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. *Journal of Economic Behavior & Organization*, *97*, 185–203.
- Englmaier, F., Kolaska, T., & Leider, S. (2016). Reciprocity in organizations—Evidence from the UK. *CESifo Economic Studies*, *62*(3), 522–546.
- Esteves-Sorenson, C. (2018). Gift exchange in the workplace: Addressing the conflicting evidence with a careful test. *Management Science*, *64*(9), 3971–4470.
- Falk, A. (2007). Gift exchange in the field. *Econometrica*, *75*(5), 1501–1511.
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, *54*(2), 293–315.
- Fehr, E., & Gächter, S. (2000). Fairness and retaliation—The economics of reciprocity. *Journal of Economic Perspectives*, *14*, 159–181.
- Fehr E., & Schmidt K. (2003). Theories of fairness and reciprocity—Evidence and economic applications. In: M. D. Dewatripont et al. (Eds.), *Advances in Economics and Econometrics, Eighth World Congress of the Econometric Society* (Vol. 1, , pp. 208–257). Cambridge: Cambridge University Press.
- Gneezy, U., & List, J. (2006). Putting behavioral economics to work: Field evidence on gift exchange. *Econometrica*, *74*(5), 1365–1384.
- Grossman, S., & Hart, O. (1983). An analysis of the principal-agent problem. *Econometrica*, *51*(1), 7–45.
- Hennig-Schmidt, H., Rockenbach, B., & Sadrieh, A. (2010). In search of workers' real effort reciprocity—A field and a laboratory experiment. *Journal of the European Economic Association*, *8*, 817–837.
- Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, *10*, 74–91.
- Ichniowski, C., & Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, *17*(1), 155–180.
- Kessler J. B. (2013). When will there be gift exchange? Addressing the lab-field debate with laboratory gift exchange experiments. Mimeo.
- Kim, M.-T., & Slonim, R. (2009). The multi-dimensional effects of reciprocity on worker effort: Evidence from a hybrid field-laboratory labor market experiment. Mimeo.
- Kube, S., Marechal, 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.
- Kube, S., Puppe, C., & Marechal, M. (2012). The currency of reciprocity—Gift exchange in the workplace. *American Economic Review*, *102*(4), 1644–1662.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*(5), 1281–1302.