

Worker reciprocity and the returns to training: evidence from a field experiment*

March 2019

Jan Sauermann[†]

Abstract

Do workers reciprocate employer-sponsored training? In this study, I use a field experiment with random assignment to training to test the hypothesis that reciprocal workers provide higher effort in response to employer-sponsored training courses compared to non-reciprocal workers. To this end, I make use of survey information on workers' reciprocal inclinations and show that reciprocal workers have higher returns to training. This result, which is robust to controlling for observed personality traits and unobserved characteristics, suggests that individuals reciprocate the firm's human capital investment with higher effort after the training, which is in line with theoretical models on gift exchange in the workplace. This finding provides an alternative rationale to explain firms' training investments even if there is risk for poaching.

JEL Codes: J24, M53, D03

Keywords: on-the-job training, reciprocity, worker performance, field experiment

*The author gratefully acknowledges financial support from the Jan Wallanders och Tom Hedelius Stiftelse for financial support (Grant number I2011-0345:1). The author thanks Sebastian Butschek, Leonie Gerhards, Christian Grund, Thomas Kittsteiner, Ilyana Kuziemko, Matthew Lindquist, Jens Mohrenweiser, Arjan Non, Gerard Pfann, and seminar participants at SOFI Stockholm, HECER Helsinki, the COPE workshop in Vienna, the Behavioral Economics Network Workshop in Stockholm, KTH Stockholm, and RWTH Aachen for valuable comments and suggestions.

[†]Swedish Institute for Social Research (SOFI), Stockholm University, 10691 Stockholm, Sweden; telephone: +46 8 16 2605; fax: +46 8 15 46 70; e-mail: jan.sauermann@sofi.su.se; Center for Corporate Performance (CCP); Institute for the Study of Labor (IZA); Research Centre for Education and the Labour Market (ROA).

1 Introduction

In labor markets challenged by globalization and technological change, training has been of interest to both policy makers and firms for a long time. In the US, 70% of employees engage in work-related training and education programs (OECD, 2017). The costs of training investments are often borne by firms, either by directly financing training courses, or indirectly by training the work force during work hours (Leuven and Oosterbeek, 1999; Caliendo et al., 2015). Why firms invest in training even if there is risk of poaching is often explained by market imperfections, such as compressed wages (Acemoglu and Pischke, 1998, 1999). An alternative argument is based on reciprocity, which can induce employees to reciprocate their employer’s investments, for example by reducing turnover, by providing greater effort, or by reducing wage demands (Leuven et al., 2005; Non, 2018). Reciprocating training investments can thereby provide a behavioral explanation for *why* firms invest in training.

The aim of this paper is to test whether positively reciprocal workers respond to training investments by exerting greater effort after participation in firm sponsored training, which should then result in higher posttraining performance. The hypothesis that reciprocity is linked to higher training returns is tested using data containing direct measures of worker performance, collected in an inhouse call center of a multinational mobile network operator in the Netherlands. In this firm, I exploit the random assignment of workers to a training program to causally estimate the returns to training (De Grip and Sauermann, 2012).¹ Arguing that the provision of this firm sponsored and fully paid week long training program can be perceived as a positive action of the firm towards the worker, I show that reciprocal workers have higher returns to the training program, that is, perform better after the training than workers with low reciprocal inclinations do after participation in the training program. This finding suggests that individuals with reciprocal attitudes return training investments by providing greater effort after participation in the training program. This is neither explained by other observed personality measures, including the Big Five and Locus of Control, nor by controlling for unobserved individual-specific effects.

¹The study of De Grip and Sauermann (2012) estimates the causal effect of training participation on performance, as well as spillover effects from trained workers to untrained peers. The personality measures used in this study have not been employed before.

Why should reciprocal inclinations matter for workplace training? Individuals with reciprocal inclinations directly receive utility from rewarding (positive reciprocity) or punishing (negative reciprocity), even if there is no direct material pay-off from this action (Rabin, 1993; Falk and Fischbacher, 2006). A large number of studies using laboratory and field experiments shows that gifts from the employer induce workers to provide greater effort than they would without the gift.² Most of these studies analyze the effect of a random treatment, that is a monetary or nonmonetary gift, on the subjects' effort provision, which allows to establish a causal link between size and type of the gift and a subject's response to the gift.

Fewer studies have used explicit measures of reciprocal attitudes derived through experiments or surveys in combination with worker and firm level outcomes. Barr and Serneels (2009) use experimentally derived measures of workers' reciprocal attitudes and measures of firm performance, and find that firms with a more reciprocal workforce are more productive than firms with a less reciprocal workforce. Cohn et al. (2015) show that performance increases after an unexpected raise in hourly wages can only be observed for workers who displayed reciprocal attitudes in a choice experiment. In environments, in which experiments are not feasible or possible, the implementation of survey questions on reciprocity in large scale surveys allows to test the importance of reciprocity for labor market outcomes (Perugini et al., 2003). Using a large representative subject pool for Germany, for example, Dohmen et al. (2009) show that workers' reciprocal attitudes are linked to higher wages, and to working longer hours.

The provision of training opportunities can be seen as one way through which firms can induce reciprocal responses. These responses could be triggered either by perceiving the training as a gift, for example if the training is of general nature and can improve outside options, or by perceiving the training as receiving the manager's attention (Dur, 2009). Indeed, Leuven et al. (2005) show that firms have an incentive to provide socially optimal levels of training if firms take workers' reciprocal inclinations into account, thereby reducing the holdup problem. Using crosssectional data for the Netherlands, Leuven et al. (2005)

²See, for example, Fehr, Gächter, and Kirchsteiger (1997), Charness (2004), Gneezy and List (2006), Hennig-Schmidt, Rockenbach, and Sadrieh (2010), Kube, Marechal, and Puppe (2012), Becker, Messer, and Wolter (2013), Cohn, Fehr, and Götte (2015), and DellaVigna et al. (2016).

provide evidence that positive reciprocal individuals are more likely to participate in employer financed training courses than individuals with low reciprocal inclinations. For Germany, Non (2018) provides corroborative evidence that reciprocal individuals are more likely to participate in training. Although the correlation could reflect unintended sorting of reciprocal individuals into training firms, the sorting pattern is also in line with the idea of strategic selection by firms to induce reciprocal incentives (Englmaier and Leider, 2012). Indeed, Englmaier et al. (2016) show that firms using personality tests when hiring are more likely to offer, amongst others, on-the-job training. Non (2018) finds that reciprocity matters in slack labor markets only, suggesting that employers use training as a means of establishing gift-exchange relationships strategically.³ In this study, randomized participation in a training program allows to estimate both the causal effect of training participation on performance, but also the indirect effect, that is the interaction between the training effect and the worker’s reciprocal attitudes. This not only helps to explain the positive correlation between training incidence and reciprocal attitudes (Leuven et al., 2005; Non, 2018), but also contributes to the general understanding of firms’ investments in employees’ human capital by offering an alternative rationale for firms’ investments in training even if there is risk of poaching (Acemoglu and Pischke, 1999; Leuven and Oosterbeek, 1999; Caliendo et al., 2015; Hoffman and Burks, 2017).

Lastly, this study also contributes to the literature on the returns to workplace training, the literature on personality in the labor market, and the role of personality in the heterogeneity in estimated returns to training. In the literature on training returns, causal interpretation of estimated returns to training has been a major issue which has mostly been addressed by fixed-effects approaches with observational data, and, to a lesser degree, with experimental variation (De Grip and Sauermann, 2012; Schwerdt et al., 2012; Adhvaryu et al., 2018). At the same time, an increasing number of studies has shown that personality traits are, amongst others, predictive for educational choices, job search, employment, and earnings. Besides reciprocity, examples are the Big Five personality traits (Bowles et al., 2001; Mueller and Plug, 2006; Gensowski, 2018), and Locus of Control (Caliendo et al., 2015,

³Although other, nonexperimental studies have explored reciprocal reactions to training courses, it is not clear to which degree these results are driven by the potential endogeneity of training participation (Mullen et al., 2006; Kampkötter and Marggraf, 2015; Montizaan et al., 2015b).

2016). With regards to training, Offerhaus (2013) finds no effect of the Big Five on training participation using a representative sample for Germany, whereas Caliendo et al. (2016) provide evidence that higher internal Locus of Control is related to higher participation rates in general training course.⁴ To the best of my knowledge, this is the first study containing both objective and repeated (panel) information on worker performance and survey evidence on personality measures, allowing to analyze how personality matters for the returns to training.

This paper is structured as follows. Section 2 provides details on the data, the field experiment, and the measures of reciprocity. The main results as well as robustness checks are presented and discussed in Section 3. Section 4 provides additional results and evidence on mechanisms. Section 5 summarizes and concludes.

2 Data and setting

2.1 Workplace, tasks, and performance measurement

The field experiment exploited in this study was organized in the inhouse call center of a multinational mobile network operator in the Netherlands from week 45/2008 to week 24/2009.⁵ This call center acts as a service center for current and prospective customers. I focus on the largest department, which only serves private customers with fixed cell phone contracts. Call agents in this department have only one task, to answer incoming customer phone calls, for example if customers have problems, complaints, or questions. Agents in this department are not involved in sales or customer acquisition. All agents take part in a training course when entering the department, which enables them to handle basic types of calls. Throughout their careers, agents regularly participate in training programs, which typically focus on learning about promotional campaigns, improving communication and IT skills, as well as learning how to handle more complex calls.

⁴In addition to these paper analyzing specific attitudes and personality, there are a number of studies analyzing cognitive and noncognitive determinants of experimental and labor market outcomes. See, for example, Bowles et al. (2001), Heckman et al. (2006), Mueller and Plug (2006) Lindqvist and Vestman (2011), Caliendo et al. (2015) and Gill and Prowse (2016).

⁵See De Grip and Sauermann (2012) for a more detailed description of data and the field experiment and the institutional background.

Call agents are organized in 10 teams, each of which is led by a team leader. During the observation period, 26 agents work for each team on average. The main purpose of being assigned to a team leader is that workers can be more efficiently supervised and monitored. There is neither team specialization, team production, nor team based incentives. Calls are typically queued before they are assigned to an available agent, irrespective of the agent's team membership. Although the firm collects large amounts of data on the performance of individual call agents, these are not explicitly used to incentivize the call agents. Agents' performance can influence wages only through an annual appraisal interview with their team leader in which agents are evaluated for the past year. Based on the outcome of this appraisal interview, agents receive an annual bonus as well as an annual wage increase. Otherwise, wages are fixed for agents.⁶

The data contain weekly information on performance outcomes, with average handling time being the most important measure for monitoring agent performance used in the firm. Average handling time is defined as the average time an agent needs to handle a customer call and is available for each individual agent and each working week.⁷ I use the inverse of average handling time multiplied by 100, which allows us to interpret high y_{it} as high performance. This performance outcome is observed for each week and each worker throughout the sample period.

2.2 The field experiment

In the department analyzed, the management introduced a new training program with the explicit aim to decrease the average time needed for handling calls. The training program was designed as a week long program, held in the call center's inhouse training center over 5 consecutive days from Monday to Friday. Call agents were paid the full time wage for the

⁶Appraisal interviews are conducted in February and March, that is before the training started. Bonuses and wage increases are paid out from July onwards, that is after the end of the observation period used in this study. Assuming that assignment to treatment and control groups does not affect the appraisal outcomes, this should not invalidate the findings presented in this paper.

⁷This measure is also used in Liu and Batt (2007), Murthy et al. (2008), and De Grip and Sauermann (2012). Agents with shorter average handling time are evaluated as performing well. The main argument for this approach is that shorter calls are cheaper for the firm. There is only limited evidence that short calls are associated with lower quality (cf. De Grip and Sauermann, 2012).

training week. Roughly half the training time was reserved for group discussions, in which the group discussed skills they lacked in their tasks, how these skills could be improved, and how the agents could provide more help to each other. During the other half of the training time, training coaches assisted the agents in handling customer calls.

The 32 week long sample period between week 45/2008 and week 24/2009 consists of three periods: a pretraining period (weeks 45/2008-09/2009), a training period (weeks 10/2009-14/2009), and a posttraining period (weeks 15/2009-24/2009). During the pretraining period, in week 50/2008, 74 out of the total 177 agents were selected for participation in the field experiment. This nonrandomly selected group was then randomly assigned to treatment and control groups.⁸ Treatment group agents were consecutively trained in small training groups during the training period; control group agents were also trained, but only after the posttraining period, which is not part of the sample period used in this paper.

Agents in the field experiment were randomly assigned to treatment and control groups by first assigning half of the teams to be trained during the training period. Due to a constraint that not more than 10 agents could be trained at once, teams were also randomly split up into separate training groups. Overall, 34 agents were trained during the training period. Control group agents ($N = 40$) were trained from week 25/2009 onwards, that is from the first week after the end of sample period used in this paper.

In this paper, I make use of survey information on reciprocity and other individual specific characteristics, which was held in the beginning of the training program. Due to partial survey nonresponse, the sample used in this study reduces to $N = 63$, consisting of 30 agents in the treatment group, and 33 agents in the control group. Column 4 of Table 1 shows that observable characteristics are balanced across treatment and control groups.⁹ Only for one of the factors of the Big Five measures, extraversion, the difference between treatment and control group is significant at the 10% significance level. Overall, this shows

⁸The firm deliberately chose to train more experienced workers to avoid losing their training investment due to high turnover among agents with low tenure. Participants in the field experiment are *not* a random sample of the overall population, but are, on average, slightly older and have longer tenure with the firm (see Table 1 for agents in the field experiment and Table A.1 in the Online appendix for the full sample of agents). This selection does not, however, violate the randomization of the field experiment, which is based on the random assignment of agents selected for the field experiment.

⁹Table A.1 in the Online Appendix further shows that there are no significant differences between individuals for the experimental sample ($N = 74$) and those who participated in the survey ($N = 63$).

that not only socioeconomic characteristics, but also personality traits are balanced across treatment and control groups, and that the treatment can be considered to be exogenous, conditional on being assigned to participate in the field experiment.

Column 1 of Table 1 show descriptive statistics for the field experimental sample. The majority of agents are women (70%), and are, on average, 36 years old. Agents have an average tenure of 4.2 years, and work as parttimers with an average of 17 weekly working hours.

2.3 Why is this setting useful to study reciprocal behavior?

There are three reasons that make this field experiment useful to study whether workers reciprocate the firm’s training investments. First, the random assignment of agents into treatment and control groups provides a setting, which allows to estimate the causal effect of the training program on performance, as well as its interaction with the measure of reciprocity. An important feature of this study is that agents in both the treatment and control groups are eventually trained and were always aware that they would eventually be trained. The estimates are identified, however, only on data *before* the control group is trained. Put differently, the effect of reciprocity is identified through within-worker variation in the timing of the training investment. Before the agents in the control group are trained, agents in the treatment group are observed for several after their training participation. This setting has two important implications for the interpretation of the results. First, although agents were neither informed about this randomization nor about the evaluation, management communicated that, due to capacity constraints in the training center, the training would be rolled out over the course of several months. For this reason, it is less likely to expect that agents in the control group perceived the treatment group’s training as unfair. Second, the fact that agents in the control group could observe or anticipate the employer’s investment even implies that the estimates are likely to be a lower bound of the true effect of the interaction between the training and reciprocity.¹⁰

¹⁰Because all agents were informed about the training in week 50/2008, it is not possible to further analyze anticipation effects.

Second, during the training program, for which the agents were paid fulltime, agents were taken out of the usual work environment and placed in a training centre separated from the work floor. Results from an internal posttraining evaluation questionnaire conducted by management suggests that agents positively experienced and enjoyed the training. Agents give an overall grade of 8.1 on a scale from 1 (worst) to 10 (best); 84% of the respondents strongly agree with the statement that “the training investment was worth the effort”.¹¹ This suggests that agents perceived the training was perceived as a positive action of the firm, rather than a chore.

Third, the theoretical motivation of this paper is that workers could perceive the training as a gift if it is of general nature. While the content of the training certainly contains some rather firm specific elements, such as knowledge related to the IT infrastructure of the firm, skills such as efficient communication with customers can be interpreted as general contents that could also be productive in other call centres, of which there are several in the direct vicinity. The effect of reciprocity should only be stronger if the training has more general contents.

2.4 Measuring reciprocity, personality, and cognitive skills

During the field experiment, call agents participated in a survey on “working in call centers,” which included statements on reciprocal attitudes, personality measures, questions to elicit cognitive test scores, and questions on socioeconomic information that is not part of the firm’s personnel data. Individual information on reciprocity was gathered using the questions developed and experimentally validated by Perugini et al. (2003), which are, amongst others, implemented in the German Socio-Economic Panel (GSOEP, Dohmen et al., 2009). In this survey, respondents were asked to rate the six questions on a 5-point Likert scale from 1 (‘does not apply to me at all’) to 5 (‘applies perfectly to me’), three of which are used to calculate a measure of positive reciprocity.¹² Although reciprocity measures derived from a

¹¹This internal survey is only partly based on individuals of the sample used in this study, and cannot be merged to the data used in this study.

¹²The questions are (1) “If someone does me a favor, I am prepared to return it,” (2) “If I suffer a serious wrong, I will take my revenge as soon as possible, no matter what the costs,” (3) “If somebody puts me in a difficult position, I will do the same to him/her,” (4) “I do my best to help somebody who helped me before,”

gift exchange experiment are more reliable (for example Cohn et al., 2015), survey evidence on reciprocity shows to yield results that are consistent with theoretical predictions (Dohmen et al., 2009; Montizaan et al., 2015a). Figure 1 shows the distribution of positive reciprocity for the estimation sample, which is similar to populationwide distributions of reciprocity (see, e.g. Dohmen et al., 2009).

In contrast to measures of reciprocity that are elicited in choice experiments, a shortcoming of survey measures is that these might reflect other individual specific characteristics. To further explore whether other individual specific characteristics can drive the findings, I employ additional measures of personality, a measure of loyalty, and test scores for cognitive skills that were all gathered in the same survey. For personality, I make use of the Big Five (conscientiousness, extraversion, agreeableness, openness to new experience, and neuroticism), Locus of Control, and negative reciprocity, each of which has been shown to have importance for labor market outcomes (Bowles et al., 2001; Mueller and Plug, 2006; Montizaan et al., 2015a; Offerhaus, 2013; Caliendo et al., 2015, 2016; Gensowski, 2018). The Big Five are measured by the 15-item questionnaire, which is implemented, amongst others, in the German Socio-Economic Panel (Gerlitz and Schupp, 2005). Each of the five factors are standardized with a mean of 0 and a standard deviation of 1. To measure Locus of Control, I employ a set of three questions, implemented from Borghans et al. (2006).¹³ As a third measure for personality, I make use of the measure of negative reciprocity, which is defined as the average of the respondent’s answers to items (2), (3), and (5) from the questions in Footnote 12.¹⁴ To measure cognitive skills, six questions on arithmetical and logical problems are implemented (CentERdata, 2007). The answers could be either right or

(5) “If somebody offends me, I will offend him/her back,” and (6) “I am ready to undergo personal costs to help somebody who helped me before.” Standardized values of the answers to items (1), (4), and (6) are used to calculate an average measure of positive reciprocal behavior. All results in this study are qualitatively similar when using measures of reciprocity, which are derived from Principal Component Analysis.

¹³The questions on Locus of Control are part of a 15-item question, which involve the concepts of Locus of Control, Anxiety, Self-image, Self-confidence, and Imagination. The questions used for Locus of Control are: “Setbacks are usually due to mistakes people make”, “Most people do not realize well to which extent their live is determined by chance”, and “Whether I reach the goals I have in my life is not a matter of luck”. All questions could be answered on a 5-point Likert scale from 1 (‘does not apply to me at all’) to 5 (‘applies perfectly to me’).

¹⁴Note that these the measures of positive and negative reciprocity are constructed such that positive and negative reciprocity are not correlated.

wrong; the cognitive test score is computed as the average of correct answers.¹⁵ The measure of loyalty is based on six questions, implemented from Mowday et al. (1979).¹⁶

3 Results

To understand the role of reciprocity in the workplace, Table 2 shows correlation coefficients between positive reciprocity on the one hand, and agent specific characteristics, personality traits, and work outcomes on the other hand. The table shows that reciprocity is not significantly related to the agents’ characteristics as measured by gender, age, and education. Among personality traits and survey measures in Column 2, only the Big Five personality measures are correlated with the measure of reciprocity, which is in line with Dohmen et al. (2008), who use representative survey data for Germany.¹⁷ The other survey based measures (cognitive test score, negative reciprocity, Locus of Control, loyalty) are not significantly correlated with reciprocity.

Among worker outcomes, only the variable leaver, which is defined as a dummy for whether an individual quits within a six months time frame after the end of the experiment, is correlated with the measure of reciprocity (Column 3 of Table 2).¹⁸ Albeit only weakly significant at the 10%-level, the interpretation of this correlation coefficient is in line with

¹⁵These questions are adapted from CentERdata (2007). The questions are (1) “There are two groups of tourists of 60 persons each. If 3/4 of the first group and 2/3 of the second group take the bus to the museum, how much larger is the first group than the second group?” (2) “A rubber ball jumps up half the distance which it fell down. If the ball falls down from a 18m high roof, how many meters does the ball travel before it touches the ground for the third time?” (3) “Four girls are 100cm, 150cm, 125cm, and 75cm, respectively. Debbie is the tallest of the four girls. Karin is the shortest. Emmy is taller than Sara. How tall is Sara?” (4) “A ball and a hat cost 1.10 Euro in total. The ball costs 1 Euro more than the hat. How much does the hat cost? Please give your answer in cents.” (5) “If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?” and (6) “In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?”. Due to partial nonresponse, the measure of cognitive skills is only available for 57 agents.

¹⁶The questions are (1) “In general, I’m telling positive things about [the firm], even if other people criticize it”, (2) “Management of [the firm] does not sufficiently motivate me to increase my performance” (negative), (3) “I’m happy that I chose [the firm] instead of a job at a different firm”, (4) “I have no interest in the development of [the firm]” (negative), and (5) “Management of [the firm] know what they do”.

¹⁷With the exception of the measure of neuroticism, the signs of the correlation coefficients between Big Five elements and the measure of reciprocity are the same as in Dohmen et al. (2008).

¹⁸Note that it is only possible to observe whether individuals left the department, but not whether individuals left the firm entirely. Because there are only limited possibilities to be promoted from call agent to management positions, it is likely that most leaving agents are leaving the firm entirely.

theory: reciprocal individuals might be more inclined to staying longer in the firm. I will discuss these outcomes in more depth in Section 4.1.

3.1 The effect of reciprocity on the returns to training

To analyze how reciprocal attitudes are related to the returns to training, the full panel structure of the field experiment (63 agents) is employed, in which performance is observed in each week an agent is working both before and after the training period from week 45/2008 to week 24/2009. Because agents are randomly assigned to participation in the training course, the causal effect of participation in the training program on worker productivity and its interaction with the measure of reciprocity can be estimated from the following ordinary least squares regression:

$$(1) \quad \log(y_{it}) = \alpha + \tau_1 d_{it} + \tau_2 rec_i + \tau_3 d_{it} \cdot rec_i + \beta_1 X_{it} + \beta_2 t_t + \beta_3 X_t + u_{it}$$

where y_{it} is the measure of productivity of worker i in week t , which is based on average handling time and for which high levels of y_{it} are interpreted as high performance. The dummy d_{it} equals one in each week after agent i 's training participation, and rec_i is the survey measure of positive reciprocity. Following De Grip and Sauermann (2012), I include control variables for several characteristics to account for remaining individual heterogeneity (X_{it}), such as working hours in week t , and whether an agent works during peak hours with high customer load in week t . To control for trends in aggregate performance and overall work load, I also include a linear time trend t_t , and the overall number of calls divided by total number of fulltime equivalent agents (X_t).¹⁹ The idiosyncratic error term u_{it} is clustered at the team level to account for the team level randomization (cf. Section 2.2 and Abadie et al. (2017)).

Table 3 shows the results of estimating Equation 1. Column 1 shows that the causal effect of training participation on performance is 0.0872, that is agents are, after participating in the training program, on average 8.7% more effective in their main task, which is answering

¹⁹Because surveys were not all conducted in the same week, fixed effects are included to control for possible survey week effects.

customer calls (cf. De Grip and Sauermann, 2012).²⁰ Column 2 shows that an agent's reciprocal attitudes are not significantly related to the outcome variable, $\log(y_{it})$, which is in line with the pairwise correlation shown in Table 2 and shows that reciprocal agents are not more productive *per se*.²¹

The coefficient of main interest in this study, $\hat{\tau}_3$, indicates whether the treatment effect is heterogenous with respect to the workers' level of reciprocity. Column 3 shows that the interaction effect between reciprocity and the randomized treatment is positive and significant: a one standard deviation difference in worker reciprocity is related to a 5.3%-point difference in the estimated returns to training. The higher treatment effect of reciprocal individuals suggests that these individuals return the training with higher effort provision after the training, possibly to return the training investment with a favorable action to their employer.²²

A potential concern is that the measure of reciprocity is correlated with unobservable characteristics, such as other personality traits or cognitive ability, which might ultimately drive what is attributed to reciprocity. An explicit way of dealing with any other characteristics would be to augment the error term in Equation 1 by an unobserved individual specific characteristic μ_i and to estimate it in a fixed effects framework, i.e. by estimating (for simplification, subscript t and additional covariates are suppressed)

$$\begin{aligned}
 \log(y_{it}) - \overline{\log(y_i)} &= \tau_1(d_{it} - \bar{d}_i) + \tau_2(rec_{it} - \overline{rec_i}) + \tau_3(d_{it} \cdot rec_{it} - \overline{d_i \cdot rec_i}) \\
 &\quad + \gamma(\mu_i - \bar{\mu}_i) + (\epsilon_i - \bar{\epsilon}_i) \\
 (2) \qquad \qquad \qquad &= \tau_1(d_{it} - \bar{d}_i) + \tau_3(d_{it} \cdot rec_{it} - \overline{d_i \cdot rec_i}) + \epsilon'_i
 \end{aligned}$$

²⁰Table A.2 in the Online Appendix shows that the reported treatment effect only slightly differs for the sample of all agents who participated in the field experiment ($N = 74$), and those who also participated in the survey ($N = 63$). This is in line with the results reported in Table A.1, which show that none of the observable characteristics differs significantly between the two groups.

²¹Note that the strong increase in the adjusted R^2 from Column 1 to 2 of Table 3 is due to the little worker-specific information Column 1 controls for.

²²An alternative interpretation of this result would be that, within an existing gift exchange between firm and agent, the provision of productivity enhancing training provides an additional margin at which agents can reciprocate the employer's actions. Because reciprocal workers are not more productive *ex ante* (cf. Table 2), this second interpretation seems to be less likely.

in which $\hat{\tau}_1$ would be the estimated (main) treatment effect. Even though the experimental design does not require individual fixed effects to estimate the causal effect of training (τ_2), fixed effects serve the purpose to control for unobserved (fixed) characteristics that are potentially correlated with the measure of reciprocity.

Column 4 of Table 3 shows the corresponding estimates for the full sample including worker fixed effects. In this regression, the estimated treatment effect of participating in the training is 2.8 percentage points lower compared to the baseline effect in Column 3 of Table 3 (11.0%). The interaction effect between the treatment and reciprocity, however, is slightly less than a third of the size in Table 3 (1.6% vs. 5.3%, cf. Column 4 of Table 3).

The results show that reciprocal individuals do have higher returns to training, despite 70% of the original effect being explained by unobserved individual specific characteristics μ_i , i.e. resulting in $Cov(\mu_i, rec_i) \neq 0$. This result also shows the importance of appropriately dealing with unobserved characteristics that might be correlated with the variable of interest.

3.2 Do other personality traits and test scores matter?

To learn more about which individual specific characteristics are possibly correlated with reciprocity, additional survey questions were gathered in the same survey as the measure of reciprocity. These include the Big Five, Locus of Control, a measure of negative reciprocity to account for personality, a measure of cognitive test scores, and a measure of loyalty. Most of these individual specific characteristics have been shown to be important drivers of economic behavior in a number of different settings. In the context of on the job training, Locus of Control has been shown to be relevant for both participation in training as well as the returns (Offerhaus, 2013; Caliendo et al., 2016).²³ As shown in Table 2 and in line with Dohmen et al. (2008), all Big Five personality elements are highly correlated with the measure of reciprocity. In contrast, the cognitive test score, Locus of Control, negative reciprocity, and the measure of loyalty are not correlated with the measure of reciprocity.

²³See, for example, Mueller and Plug (2006), Lindqvist and Vestman (2011) and Gensowski (2018) for the importance of personality, Lindqvist and Vestman (2011) and Gensowski (2018) for cognitive skills, and Dohmen et al. (2009) and Montizaan et al. (2015a) for negative reciprocity.

To test whether these characteristics may drive the coefficient of main interest, that is the interaction effect between reciprocity and the treatment ($\hat{\tau}_3$), Equation 1 is estimated by separately including each of the individual specific characteristics. For the Big Five elements, Columns 1 to 6 of Table 4 show that neither of the Big Five elements nor all five elements altogether significantly affect performance. Columns 7 to 10 show the corresponding estimates when including measures of the cognitive test score, negative reciprocity, Locus of Control, and loyalty, respectively. Although none of these variables are significantly correlated with the measure of positive reciprocity, one could argue that these characteristics could pick up similar treatment effect heterogeneity. Only the measure of loyalty is significantly related to the outcome variable. The coefficient of interest, the interaction between reciprocity and the treatment, always remains significant and stable throughout most regressions (0.046-0.061).²⁴

4 Additional results

4.1 Alternative channels of reciprocal behavior

The results so far have provided evidence for treatment effect heterogeneity with respect to workers' degree of positive reciprocity. Although the descriptive analysis provided only limited evidence for significant correlation between reciprocity and other worker outcomes, reciprocal individuals might have alternative, possibly competing, channels through which training investments are returned. Examples for these competing channels are the number of hours worked by the agent, the number of hours of absence, for example due to sickness, and turnover and fluctuation.²⁵

²⁴Table A.3 in the Online Appendix shows corresponding estimation results when also interacting the added characteristics with the treatment. Only when including the Big Five element of neuroticism and its interaction effect in the regression, or when including all five elements in one regression, the interaction effect between reciprocity and the treatment becomes insignificant (with p values of 0.134 and 0.120, respectively). Without affecting the interaction between reciprocity and the treatment, the table also shows that individuals with higher levels of loyalty, extraversion, and higher cognitive test scores have higher returns to training.

²⁵There are two alternative arguments, which cannot be tested with this setting. First, Leuven et al. (2005) put forward the argument that employees could reciprocate training with lower wage demands. In the context of this field experiment, however, this hypothesis is not testable, because there is no information on wage bargaining. Second, reciprocal agents might have higher returns to training because they perceive the training in a way that the firm pays attention to them, or even interpret the training investment as a

First, reciprocal individuals could return the training investments by working longer hours. For the firm, this would result in a better utilization of labor. Column 2 of Table 5 shows that when estimating Equation 1 with the number of working hours as an outcome variable, the interaction variable is not significant. When using the preferred specification including individual fixed-effects to control for unobserved individual-specific characteristics that might be correlated with the measure of reciprocity, the estimate is significant and positive. This result suggests that workers with a one standard deviation higher reciprocity increase their working time by 0.4 hours per week (Column 3).

Second, reciprocity might have a negative effect on hours of absence. If reciprocal individuals would have lower degrees of sickness absence after the training, this would suggest that the training investment is reciprocated with lower absence. Kampkötter and Marggraf (2015) and Adhvaryu et al. (2018) have previously shown that training participation can affect these outcomes. To address this channel, Equation 1 is estimated with an outcome variable *hours of absence*. Column 5 of Table 5 shows that the estimated coefficient $\hat{\tau}_3$ has the expected negative sign and is significantly different from zero. When including individual fixed effects, however, this effect vanishes, likely because this is driven by other individual-specific characteristics (Column 6).

Third, an alternative way to reciprocate the training investments would be decreased labor turnover, i.e. that workers are less likely to leave the firm. Table 2 shows that the measure of reciprocity is significantly correlated with a dummy variable whether or not an employee left the department within six months after participating in the survey: reciprocal individuals are less likely to have left within this period. Because *all* agents are trained eventually in this setting, it is not possible to estimate the effect on this measure. Instead, Columns 7 and 8 of Table 5 use a measure defined as leaving the department before the end of the observation period, i.e. before the control group gets trained. The results show that neither participation in the training program itself, nor the interaction with reciprocity significantly affected the probability to leave the department. This result, however, should be treated with caution since only four agents have left the department in the weeks between

signal that they would not be laid off in the near future. These arguments could be tested in other settings, in which the experimenter could, for instance, randomly vary layoff probabilities.

the training period and the end of the observation period, and also because the definition of the dependent variable does not allow to control for individual fixed effects, which has been shown to be important in previous regressions.

4.2 Are agents returning the favor of being selected?

The results shown in Section 3 are in line with an interpretation of reciprocal inclinations generating higher returns to training. An alternative strategic motive is that agents selected to be part of the training program feel that they are part of an exclusive group and therefore feel more motivated.²⁶ This motive, which might be correlated with reciprocity, could trigger similar reactions and might explain why the fixed-effects estimates reduce the interaction effect between reciprocity and the treatment.

Agents in both treatment and control groups were selected to be part of the training program, but were then randomly assigned to either group. That is, the effects in this paper are identified from exogenous variation in participating in the training course, conditional on being part of the training program. Due to the randomization in the experiment, tenure, which was the main criteria of being part of the training program, is balanced across treatment and control groups. To shed more light on whether tenure affects the interaction effect between the treatment and reciprocity, Table A.4 estimates Equation 1 augmented by agents' tenure and an interaction effect between the treatment and tenure. The estimates show that neither tenure nor its interaction effect with the treatment dummy has a significant effect, whereas the interaction effect with reciprocity remains unchanged. Although it is not possible to rule out that being selected into the training program *did* affect motivation and created increased performance, it does not drive the interaction effect between the treatment and reciprocity.

4.3 Reciprocating during or after the training?

An important question is whether reciprocal individuals put more effort into their training, that is whether it is that reciprocal individuals are better in utilizing the training, or are they

²⁶I thank an anonymous referee for highlighting this point.

exerting greater effort *after* participation in training. Whereas the former, more efficient human capital acquisition, should lead to a more permanent increase in skills and thus performance, the latter might cause only a transitory effect on performance that fades out.

Figure 2 shows the treatment effect and the interaction with reciprocity, separately for each week after the training. The black line shows that the treatment effect of participation in the training program reaches a peak in the fifth week after training, and decreases substantially afterwards, which could be explained by spill-over effects (De Grip and Sauer-mann, 2012), motivational effects, or even human capital depreciation. The solid gray line shows the interaction effect between the treatment dummy and reciprocity by week after the training and thereby corresponds to the interaction effect shown in Column 4 of Table 3. Despite being small in size, this effect does not follow this decrease (gray line). Towards the end of the observation period (week 11 after the training period), the point estimate of the interaction effect has almost the same magnitude as the main effect, suggesting that reciprocal individuals have a small yet more permanent effect than nonreciprocal individuals. Although far from conclusive, this could be explained by higher effort provision during the training, resulting in more efficient human capital acquisition and more stable increases in performance.

4.4 Measurement error vs. unobserved characteristics

Why is the interaction effect of reciprocity and the treatment dummy so much smaller when controlling for worker-specific fixed effects? To shed more light on this issue, I split the sample in individuals with below median and above median levels of reciprocity, which allows to compare the results with and without fixed effects. Columns 1 and 2 of Table 6 show that, for individuals with low levels of reciprocity, the point estimate of the returns to the training program are very similar, albeit more precisely estimated when controlling for individual fixed effects. For individuals with high levels of reciprocity, Column 3 shows that participation in the training program leads to a 15.5% higher performance after the training. Including individual fixed effects reduces this estimate by 5%-points to 10.6%. There are two explanations for this diverging pattern of individuals with low levels and high levels

of reciprocity. First, one could argue that the sample of treatment and control groups are not perfectly balanced for the sample of individuals with high levels of reciprocity. This could lead to a biased estimate of the returns to training, which could then be reduced by including individual fixed effects. Replicating the t -test from Table 1 for the sample of individuals with high levels of reciprocity ($N = 32$), however, does not show any significant differences (Table A.5). Second, unobserved individual specific characteristics could matter more for individuals with high levels of reciprocity, than they do for individuals with low levels of reciprocity. This could also be the case, for example, if reciprocity is measured with measurement error in the lower part of the reciprocity distribution.

4.5 Negative effects on non-trained workers

Given that the training was organized as a week long training program, an important question is how removing up to 10 workers in one week from the work floor affects the remaining agents.²⁷ This implies that during training weeks the headcount in training weeks is up to 10% lower at the department level, and up to 66% at the team level.

To test whether removing agents from teams for training purposes affects the remaining workers, Table A.6 tests whether the share of co-workers in training affects contemporaneous performance of the remaining agents. It shows that having a higher share of co-workers participating in the training courses poses a modest negative effect on an agents performance in the week that the co-workers are trained. This effect, however, is imprecisely estimated.

5 Conclusions

Although there is empirical evidence that workers with positive reciprocal inclinations participate more often in firm sponsored training (Leuven et al., 2005), there is no evidence on whether and how individuals reciprocate the firm's training investments. Despite using data from a field experiment in a call centre in the Netherlands, this study is able to provide previously undocumented evidence that individuals reciprocate firm sponsored training. To test

²⁷It should be noted that all agents in the department do the same task, i.e. calls are randomly allocated to agents available in the department.

this hypothesis, personnel data with panel information on worker performance is combined with random assignment to training courses, and direct measures of reciprocal attitudes. Compared to survey datasets, this setting allows to causally identify treatment effects of participation in the training program, as well as its interaction with survey measures of reciprocity. Controlling for time-invariant individual-specific characteristics, the results show that individuals with reciprocal inclinations have higher returns to training. This suggests that workers reciprocate the firm’s training investments by greater effort provision, thereby translating into higher performance.

The finding that individuals with reciprocal inclinations have higher returns to training has important implications for the worker firm relationship: first, the finding provides an alternative channel through which the gift exchange between workers and firms can operate. Second, it suggests an alternative argument for why firms invest in human capital, even if there is risk of poaching. Workers’ reciprocal attitudes increase the incentive for firms to invest in training, and to positively select workers in the hiring process (cf. Englmaier et al., 2016). Third, the result individual fixed effects substantially reduce the estimated effect of interest show that it is crucial to consider alternative personality measures and individual fixed-effects to account for worker-specific heterogeneity that might otherwise drive effects.

References

- Abadie, A., Athey, S., Imbens, G., Wooldridge, J., 2017. When should you adjust standard errors for clustering? NBER Working Paper 24003, National Bureau of Economic Research (NBER).
- Acemoglu, D., Pischke, J.-S., 1998. Why do firms train? theory and evidence. *Quarterly Journal of Economics* 113 (1), 79–119.
- Acemoglu, D., Pischke, J.-S., 1999. The structure of wages and investment in general training. *Journal of Political Economy* 107, 539–72.

- Adhvaryu, A., Kala, N., Nyshadham, A., February 2018. The skills to pay the bills: Returns to on-the-job soft skills training. Working Paper 24313, National Bureau of Economic Research.
- Barr, A., Serneels, P., 2009. Reciprocity in the workplace. *Experimental Economics* 12 (1), 99–112.
- Becker, S. O., Messer, D., Wolter, S. C., 2013. A gift is not always a gift: Heterogeneity and long-term effects in a gift exchange experiment. *Economica* 80, 345–71.
- Borghans, L., Golsteyn, B., De Grip, A., March 2006. Meer werken is meer leren: Determinanten van kennisontwikkeling. Report, CINOP, 's-Hertogenbosch.
- Bowles, S., Gintis, H., Osborne, M., December 2001. The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39 (4), 1137–76.
- Caliendo, M., Cobb-Clark, D. A., Seitz, H., Uhlenhorff, A., 2016. Locus of control and investment in training. IZA Discussion Paper 10406, Institute of Labor Economics (IZA).
- Caliendo, M., Cobb-Clark, D. A., Uhlenhorff, A., March 2015. Locus of Control and Job Search Strategies. *Review of Economics and Statistics* 97 (1), 88–103.
- CentERdata, 2007. Codeboek lerend individu 2007. Codebook, Tilburg University, Tilburg.
- Charness, G., 2004. Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics* 22 (3), 665–88.
- Cohn, A., Fehr, E., Goette, L., 2015. Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science* 61 (8), 1777–94.
- De Grip, A., Sauermann, J., 2012. The effects of training on own and co-worker productivity: Evidence from a field experiment. *Economic Journal* 122 (560), 376–99.
- DellaVigna, S., List, J. A., Malmendier, U., Rao, G., February 2016. Estimating social preferences and gift exchange at work. Working Paper 22043, National Bureau of Economic Research.

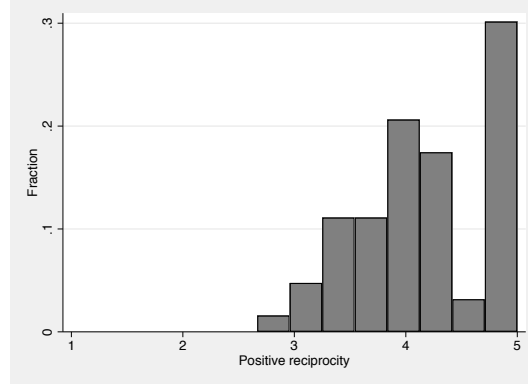
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2008. Representative trust and reciprocity: Prevalence and determinants. *Economic Inquiry* 46 (1), 84–90.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2009. Homo reciprocans: Survey evidence on behavioral outcomes. *Economic Journal* 119 (536), 592–612.
- Dur, R., 2009. Gift exchange in the workplace: Money or attention? *Journal of the European Economic Association* 7 (23), 550–60.
- Englmaier, F., Kolaska, T., Leider, S., 2016. Reciprocity in Organizations: Evidence from the UK. *CESifo Economic Studies* 62 (3), 522–46.
- Englmaier, F., Leider, S., May 2012. Contractual and organizational structure with reciprocal agents. *American Economic Journal: Microeconomics* 4 (2), 146–83.
- Falk, A., Fischbacher, U., February 2006. A theory of reciprocity. *Games and Economic Behavior* 54 (2), 293–315.
- Fehr, E., Gächter, S., Kirchsteiger, G., 1997. Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica* 65 (4), 833–60.
- Gensowski, M., 2018. Personality, IQ, and Lifetime Earnings. *Labour Economics* 51, 170–83.
- Gerlitz, J.-Y., Schupp, J., 2005. Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. *Research Notes* 4, German Institute for Economic Research (DIW), Berlin.
- Gill, D., Prowse, V., 2016. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy* 124 (6), 1619–1676.
- Gneezy, U., List, J. A., 2006. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica* 74 (5), 1365–84.
- Heckman, J. J., Stixrud, J., Urzua, S., July 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics* 24 (3), 411–82.

- 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 (4), 817–37.
- Hoffman, M., Burks, S. V., March 2017. Training contracts, employee turnover, and the returns from firm-sponsored general training. Working Paper 23247, National Bureau of Economic Research.
- Kampkötter, P., Marggraf, K., 2015. Do employees reciprocate to intra-firm trainings? An analysis of absenteeism and turnover rates. *International Journal of Human Resource Management* 26 (22), 2888–907.
- Kube, S., Maréchal, M. A., Puppe, C., 2012. The currency of reciprocity: Gift exchange in the workplace. *American Economic Review* 102 (4), 1644–62.
- Leuven, E., Oosterbeek, H., 1999. Demand and supply of work-related training: Evidence from four countries. In: *Research in Labor Economics*. 18, pp. 303–30.
- Leuven, E., Oosterbeek, H., Sloof, R., Van Klaveren, C., 2005. Worker reciprocity and employer investment in training. *Economica* 72 (285), 137–49.
- Lindqvist, E., Vestman, R., January 2011. The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics* 3 (1), 101–28.
- Liu, X., Batt, R., 2007. The economic pay-offs to informal training: Evidence from routine service work. *Industrial and Labor Relations Review* 61 (1), 75–89.
- Montizaan, R., Cörvers, F., De Grip, A., Dohmen, T., 2015a. The impact of negatively reciprocal inclinations on worker behavior: Evidence from a retrenchment of pension rights. *Management Science* 62 (3), 668–81.
- Montizaan, R., de Grip, A., Fouarge, D., 2015b. Training access, reciprocity, and expected retirement age. IZA Discussion Paper 8862, Institute of Labor Economics (IZA).

- Mowday, R. T., Steers, R. M., Porter, L. W., 1979. The measurement of organizational commitment. *Journal of Vocational Behavior* 14 (2), 224–47.
- Mueller, G., Plug, E., 2006. Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review* 60 (1), 3–22.
- Mullen, T. R., Kroustalis, C., Meade, A. W., Surface, E. A., 2006. Assessing change in perceived organizational support due to training. Conference paper, Paper presented at the 21st annual meeting of the Society for Industrial/Organizational Psychology, Dallas, TX.
- Murthy, N. N., Challagalla, G. N., Vincent, L. H., Shervani, T. A., 2008. The impact of simulation training on call center agent performance: A field-based investigation. *Management Science* 54 (2), 384–99.
- Non, A., 2018. Training participation and the role of reciprocal attitudes. ROA Research Memorandum 005, Research Centre for Education and the Labour Market (ROA).
- OECD, 2017. *OECD Skills Outlook 2017: Skills and Global Value Chains*. OECD Publishing, Paris.
- Offerhaus, J., 2013. The type to train? impacts of personality characteristics on further training participation. *SOEPpapers on Multidisciplinary Panel Data Research* 531, DIW Berlin, The German Socio-Economic Panel (SOEP), Berlin.
- Perugini, M., Gallucci, M., Presaghi, F., Ercolani, A. P., 2003. The personal norm of reciprocity. *European Journal of Personality* 17, 251–83.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83 (5), 1281–302.
- Schwerdt, G., Messer, D., Woessmann, L., Wolter, S. C., 2012. The impact of an adult education voucher program: Evidence from a randomized field experiment. *Journal of Public Economics* 96 (7-8), 569–583.

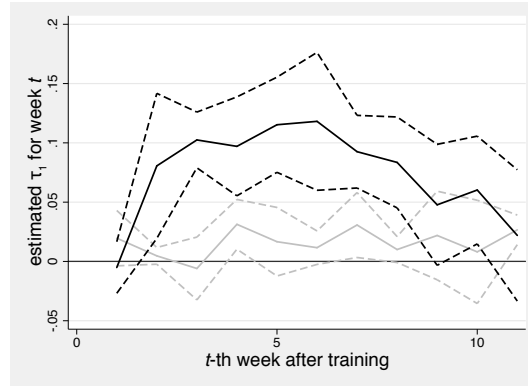
Figures

Figure 1: Distribution of reciprocity



Note: The figure shows the histogram of positive reciprocity for the estimation sample ($N = 63$). The underlying questions (see Section 2.4) could be answered on a scale from 1 ('does not apply to me at all') to 5 ('applies perfectly to me').

Figure 2: Treatment and interaction effects on performance over time



Note: This figure shows the estimated treatment effect on performance for each week after the training (solid black line) and the corresponding 95% confidence interval (dashed black lines). The gray lines shows the estimated interaction effect between treatment dummy and the measure of reciprocity (solid gray line) and the corresponding 95% confidence intervals (dashed gray lines). The estimates are based on a regression controlling for individual fixed effects, working hours, the share of peak hours, calls per FTE, a linear time trend, and for the week the survey was taken (cf. the regression shown in Column 4 of Table 3). Week 0 denotes the training week.

Tables

Table 1: Descriptive statistics

Sample	(1) All	(2) Treat. group	(3) Control group	(4) Diff (3)–(2)
Gender (1=male)	0.3016 (0.4626)	0.3667 (0.4901)	0.2424 (0.4352)	-0.1242 (-1.0658)
Age	36.3504 (11.2356)	34.9622 (10.3689)	37.6125 (11.9876)	2.6503 (0.9341)
Tenure (in years)	4.1990 (3.9701)	4.4147 (3.9226)	4.0029 (4.0633)	-0.4118 (-0.4084)
High education	0.3103 (0.4667)	0.3571 (0.4880)	0.2667 (0.4498)	-0.0905 (-0.7348)
Performance	0.3629 (0.0837)	0.3673 (0.0727)	0.3589 (0.0935)	-0.0085 (-0.3985)
Number of calls	196.4127 (115.7966)	187.0000 (119.1478)	204.9697 (113.8235)	17.9697 (0.6121)
Working hours	16.6508 (8.4799)	15.9667 (9.3199)	17.2727 (7.7309)	1.3061 (0.6074)
Share Peak-Hours	0.5328 (0.1935)	0.5386 (0.1817)	0.5276 (0.2063)	-0.0110 (-0.2242)
Absenteeism	0.1111 (0.3168)	0.1000 (0.3051)	0.1212 (0.3314)	0.0212 (0.2634)
Training incidence	0.1905 (0.3958)	0.1000 (0.3051)	0.2727 (0.4523)	0.1727 (1.7588)
Leaver	0.5238 (0.5034)	0.5667 (0.5040)	0.4848 (0.5075)	-0.0818 (-0.6412)
Positive reciprocity	4.2011 (0.6627)	4.1222 (0.7349)	4.2727 (0.5919)	0.1505 (0.8989)
Positive reciprocity (std.)	0.0118 (1.0283)	-0.1105 (1.1403)	0.1230 (0.9184)	0.2335 (0.8989)
Conscientiousness	12.6667 (1.4142)	12.5667 (1.3566)	12.7576 (1.4797)	0.1909 (0.5320)
Extraversion	12.0952 (1.8554)	12.6667 (1.5388)	11.5758 (1.9848)	-1.0909* (-2.4204)
Agreeableness	12.7619 (1.6821)	12.5667 (1.8696)	12.9394 (1.4987)	0.3727 (0.8767)
Openness to experience	10.3810 (1.9380)	10.5333 (1.9250)	10.2424 (1.9690)	-0.2909 (-0.5919)
Neuroticism	7.2857 (2.3721)	7.2667 (2.4344)	7.3030 (2.3517)	0.0364 (0.0603)
Cognitive test score	0.4561 (0.2450)	0.4770 (0.2736)	0.4345 (0.2144)	-0.0425 (-0.6511)
Negative reciprocity	2.4815 (0.8875)	2.5556 (0.8502)	2.4141 (0.9281)	-0.1414 (-0.6285)
Locus of Control	8.7143 (1.5390)	8.8667 (1.6132)	8.5758 (1.4797)	-0.2909 (-0.7466)
Loyalty	1.1270 (0.7042)	1.1467 (0.6453)	1.1091 (0.7634)	-0.0376 (-0.2099)
Observations	63	30	33	63

Note: Standard deviations are in parentheses in Columns 1 to 3; *t*-statistics in Column 4.

Table 2: Correlates of positive reciprocity

(1) Worker characteristics		(2) Personality measures		(3) Worker outcomes	
Gender (1=male)	0.1104	<i>Big Five</i>		Performance	-0.0295
Age	-0.0124	Conscientiousness	0.4020***	Tenure	0.1515
High skilled education	0.0929	Extraversion	0.0898	Working hours	0.0109
		Agreeableness	0.2404*	Share peak hours	0.0192
		Openness	0.2487**	Absenteeism	-0.0841
		Neuroticism	-0.2997**	Training incidence	-0.0466
		Cognitive test score	0.1907	Leaver	-0.2127*
		Negative reciprocity	-0.0115		
		Locus of Control	-0.0763		
		Loyalty	-0.1450		

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Figures in this table show the pairwise Pearson's correlation coefficients of workers' characteristics with the measure of positive reciprocity (standardized). The figures are based on the estimation sample and contains one observation per agent ($N = 63$, cf. Column 1 of Table 1). All personality measures are standardized with 0 mean and standard deviation of 1. All time varying variables in Column 3 are averaged over the time between the start of the sample (45/2008) and the start of the field experiment (09/2009). Worker outcomes are defined as: performance (inverse of average handling time), tenure in years, weekly working hours, share of working hours during peak hours, absenteeism (share of weeks an agent reported being sick), training incidence (share of weeks in which an agent received training), leaver is defined whether a worker left within 6 months after the end of the experiment.

Table 3: The returns to training and interaction with reciprocity

	(1)	(2)	(3)	(4)
Treatment dummy	0.0872*** (0.0227)	0.1063*** (0.0161)	0.1104*** (0.0153)	0.0860*** (0.0133)
Reciprocity (std.)		0.0124 (0.0249)	0.0015 (0.0262)	
Treatment \times reciprocity			0.0527** (0.0219)	0.0158** (0.0057)
Calls per FTE	0.0001 (0.0001)	0.0001** (0.0001)	0.0001** (0.0001)	0.0001** (0.0000)
Working hours	0.0018 (0.0022)	0.0010 (0.0012)	0.0011 (0.0012)	-0.0017 (0.0011)
Share Peak-Hours	-0.2736*** (0.0813)	-0.2784*** (0.0658)	-0.2872*** (0.0596)	-0.3592*** (0.0390)
Time trend	0.0020 (0.0013)	0.0015* (0.0008)	0.0014* (0.0008)	0.0024*** (0.0005)
Constant	-1.0833*** (0.1075)	-0.9946*** (0.0702)	-0.9875*** (0.0661)	-0.9764*** (0.0628)
Observations	1,672	1,672	1,672	1,672
Number of agents	63	63	63	63
Adjusted R-squared	0.0740	0.2718	0.2799	0.6127
Worker FE	No	No	No	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. All regressions also include fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level.

Table 4: The effect of reciprocity and worker specific characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment dummy	0.1112*** (0.0157)	0.1099*** (0.0147)	0.1131*** (0.0147)	0.1117*** (0.0169)	0.1102*** (0.0141)	0.1182*** (0.0159)	0.1137*** (0.0157)	0.1155*** (0.0156)	0.1114*** (0.0147)	0.1118*** (0.0193)
Reciprocity (std.)	0.0043 (0.0253)	0.0018 (0.0269)	0.0004 (0.0290)	0.0081 (0.0279)	-0.0073 (0.0272)	-0.0008 (0.0313)	0.0056 (0.0250)	0.0022 (0.0268)	0.0027 (0.0278)	0.0070 (0.0208)
Treatment \times reciprocity	0.0544** (0.0234)	0.0545** (0.0218)	0.0537** (0.0217)	0.0500** (0.0223)	0.0512** (0.0207)	0.0451** (0.0185)	0.0612** (0.0243)	0.0560** (0.0223)	0.0539** (0.0225)	0.0460** (0.0185)
Conscientiousness	-0.0058 (0.0262)					-0.0042 (0.0285)				
Extraversion		0.0033 (0.0110)				-0.0151 (0.0097)				
Agreeableness			0.0082 (0.0150)			0.0053 (0.0176)				
Openness				-0.0432 (0.0307)		-0.0374 (0.0314)				
Neuroticism					-0.0295 (0.0174)	-0.0318 (0.0206)				
Cognitive test score							-0.0742 (0.0533)			
Negative reciprocity								-0.0161 (0.0163)		
Locus of Control									0.0052 (0.0172)	
Loyalty										0.0543** (0.0184)
Constant	-0.9818*** (0.0723)	-0.9926*** (0.0713)	-0.9853*** (0.0651)	-0.9955*** (0.0821)	-0.9991*** (0.0764)	-0.9791*** (0.0700)	-0.9334*** (0.0714)	-0.9759*** (0.0662)	-0.9833*** (0.0799)	-0.9509*** (0.0693)
Observations	1,672	1,672	1,672	1,672	1,672	1,672	1,531	1,672	1,672	1,672
Number of agents	63	63	63	63	63	63	57	63	63	63
Adjusted R-squared	0.2796	0.2797	0.2805	0.2987	0.2939	0.3098	0.2247	0.2835	0.2799	0.3236

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. All regressions also include a constant, working hours, share in peak hours, calls per fulltime equivalents (FTE), a linear time trend, and fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level. The F-Statistic and corresponding p -value corresponds to a t -test of the main and interaction effects of all added personality traits and cognitive test scores. Due to partial nonresponse, the sample used in Column 7 is based on 57 agents instead of 63 agents.

Table 5: The returns to training on working hours, sickness absence, and turnover

Outcome variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Working hours			Hours of absence			Turnover	
Treatment dummy	-1.1631*	-1.2477*	-0.9891	0.0180***	0.0158**	-0.0142	-0.0072	-0.0079
	(0.6341)	(0.6189)	(0.9021)	(0.0051)	(0.0060)	(0.0172)	(0.0811)	(0.0803)
Reciprocity (std.)	-0.0860	0.1557		-0.0153	-0.0096		0.0073	0.0094
	(0.9537)	(1.0073)		(0.0142)	(0.0157)		(0.0339)	(0.0413)
Treatment \times reciprocity		-1.1609	0.4089**		-0.0275*	-0.0126		-0.0099
		(0.9827)	(0.1745)		(0.0140)	(0.0178)		(0.0406)
Constant	22.0926***	21.8515***	22.9772***	0.2777**	0.2740**	0.3922***	0.5060	0.5047
	(3.7003)	(3.6349)	(4.2591)	(0.0892)	(0.0877)	(0.0704)	(0.3005)	(0.3034)
Observations	1,672	1,672	1,672	1,672	1,672	1,672	1,672	1,672
Number of agents	63	63	63	63	63	63	63	63
Adjusted R-squared	0.2095	0.2121	0.5059	0.1120	0.1143	0.2471	0.2023	0.2020
Individual FE	No	No	Yes	No	No	Yes	No	No

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: number of working hours (Columns 1 to 3), hours of absence (Columns 4 to 6), and leaving the department (Columns 7 and 8). All regressions also include calls per FTE, a linear time trend, the number of working hours, the share of peak hours, and fixed effects for the week in which agents participated in the survey. Because turnover is time-invariant, fixed-effect estimation is not included. Standard errors are clustered at the team level.

Table 6: The returns to training by low and high reciprocal individuals

Sample	(1)	(2)	(3)	(4)
	Below median reciprocity		Above median reciprocity	
Treatment dummy	0.0632	0.0608**	0.1551***	0.1062***
	(0.0483)	(0.0208)	(0.0356)	(0.0201)
Constant	-0.9671***	-0.9129***	-0.9298***	-1.0869***
	(0.1165)	(0.0825)	(0.1804)	(0.1057)
Observations	846	846	826	826
Number of agents	31	31	32	32
(TG / CG)	(16/15)	(16/15)	(14/18)	(14/18)
Adjusted R-squared	0.2214	0.6329	0.3841	0.5948
Individual FE	No	Yes	No	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. All regressions also include working hours, share in peak hours, calls per fulltime equivalents (FTE), a linear time trend, and fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level.

Online appendix to

“Worker reciprocity and the returns to training: evidence from a field experiment”

Jan Sauermann (Friday 1st March, 2019)

Table A.1: Descriptive statistics of agents in the call centre and the field experiment

Sample	(1)	(2)	(3)	(4)
	All agents	All	Field experiment With survey	Diff (2)–(3)
Gender (1=male)	0.2903 (0.4554)	0.3243 (0.4713)	0.3016 (0.4626)	0.1074 (0.9761)
Age	32.9498 (11.3471)	35.1918 (10.9330)	36.3504 (11.2356)	-2.0761 (-0.8122)
Tenure (in years)	2.7784 (3.6260)	3.9101 (3.8813)	4.1990 (3.9701)	0.2982 (0.3273)
Performance	0.3413 (0.1035)	0.3588 (0.0837)	0.3629 (0.0837)	0.0013 (0.0664)
Number of calls	200.9484 (99.8692)	194.0541 (111.9994)	196.4127 (115.7966)	-23.6059 (-0.9024)
Working hours	18.6516 (8.0078)	16.5946 (8.1508)	16.6508 (8.4799)	-1.6441 (-0.8632)
Share peak hours	0.5545 (0.1710)	0.5409 (0.1941)	0.5328 (0.1935)	0.0266 (0.5853)
Absenteeism	0.0710 (0.2576)	0.1216 (0.3291)	0.1111 (0.3168)	-0.0074 (-0.0951)
Training incidence	0.2065 (0.4061)	0.1892 (0.3943)	0.1905 (0.3958)	-0.1324 (-1.4497)
Leaver	0.6387 (0.4819)	0.5676 (0.4988)	0.5238 (0.5034)	0.0382 (0.3266)
Observations	155	74	63	

Note: The sample in Column 1 is defined as all agents working in the department during the observation period. Column 2 contains all agents who participated in the field experiment. Column 3 contains all agents who participated in the field experiment and who participated in the survey. Column 4 shows the difference between agents with survey participation and those without. Parentheses in Columns 1 to 3 contain standard deviations; parentheses in Column 4 contain *t*-statistics.

Table A.2: The returns to training and sample definition

Sample	(1) Initial randomization	(2) Initial randomization	(3) Estimation sample	(4) Estimation sample
Treatment dummy	0.0836*** (0.0217)	0.0882*** (0.0206)	0.0872*** (0.0227)	0.0860*** (0.0209)
Calls per FTE	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Working hours	0.0015 (0.0022)	-0.0015 (0.0010)	0.0018 (0.0022)	-0.0017 (0.0011)
Share Peak-Hours	-0.2478*** (0.0588)	-0.3416*** (0.0749)	-0.2736*** (0.0813)	-0.3517*** (0.0794)
Time trend	0.0024 (0.0013)	0.0023** (0.0009)	0.0020 (0.0013)	0.0024*** (0.0009)
Constant	-1.1281*** (0.1027)	-1.0000*** (0.0888)	-1.0833*** (0.1075)	-0.9827*** (0.0931)
Observations	1,858	1,859	1,672	1,673
Number of agents	74	74	63	63
Adjusted R-squared	0.0739	0.5997	0.0740	0.6119
Individual FE	No	Yes	No	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. The sample in Columns 1 and 2 is defined as all agents who participated in the field experiment ($N = 74$, cf. De Grip and Sauermann, 2012). The sample in Columns 3 and 4 is the estimation sample used in this study ($N = 63$), and is defined as agents from the field experiment who also participated in the survey. All regressions also include fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level.

Table A.3: Reciprocity and other worker specific characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment dummy	0.1085*** (0.0165)	0.0854*** (0.0110)	0.1099*** (0.0150)	0.1127*** (0.0196)	0.1085*** (0.0136)	0.0996*** (0.0170)	0.0429 (0.0261)	0.1176*** (0.0186)	0.1112*** (0.0172)	0.1107*** (0.0193)
Reciprocity	0.0037 (0.0238)	0.0011 (0.0260)	0.0005 (0.0284)	0.0065 (0.0271)	-0.0073 (0.0264)	-0.0037 (0.0283)	0.0059 (0.0243)	0.0017 (0.0264)	0.0019 (0.0271)	0.0054 (0.0217)
Treatment × reciprocity	0.0488** (0.0168)	0.0543** (0.0211)	0.0486* (0.0220)	0.0525* (0.0245)	0.0420 (0.0266)	0.0557 (0.0335)	0.0462* (0.0215)	0.0562** (0.0225)	0.0525** (0.0222)	0.0424** (0.0183)
Conscientiousness	-0.0068 (0.0290)					0.0018 (0.0328)				
Treatment × conscientiousness	0.0125 (0.0259)					-0.0188 (0.0306)				
Extraversion		-0.0021 (0.0101)				-0.0207** (0.0081)				
Treatment × extraversion		0.0553** (0.0232)				0.0492*** (0.0142)				
Agreeableness			0.0049 (0.0147)			0.0015 (0.0183)				
Treatment × agreeableness			0.0163 (0.0148)			0.0120 (0.0132)				
Openness to experience				-0.0392 (0.0290)		-0.0331 (0.0284)				
Treatment × openness				-0.0213 (0.0515)		-0.0177 (0.0536)				
Neuroticism					-0.0266* (0.0143)	-0.0332* (0.0160)				
Treatment × neuroticism					-0.0187 (0.0319)	0.0101 (0.0330)				
Cognitive test score							-0.0978* (0.0539)			
Treatment × cog. test							0.1549*** (0.0344)			
Negative reciprocity								-0.0121 (0.0095)		
Treatment × neg. reciprocity								-0.0215 (0.0316)		
Standardized values of (pers_loc)									0.0064 (0.0183)	
Treatment × Locus of Control									-0.0053 (0.0205)	
Loyalty										0.0475** (0.0181)
Treatment × loyalty										0.0395* (0.0213)
Constant	-0.9828*** (0.0715)	-0.9825*** (0.0690)	-0.9817*** (0.0653)	-0.9987*** (0.0794)	-0.9955*** (0.0726)	-0.9739*** (0.0682)	-0.9225*** (0.0687)	-0.9739*** (0.0663)	-0.9825*** (0.0796)	-0.9492*** (0.0687)
F-stat controls=0	0.163 (0.852)	2.962 (0.0977)	0.653 (0.541)	1.055 (0.384)	1.737 (0.225)	680788 (0)	13.77 (0.00134)	1.163 (0.351)	0.0711 (0.932)	43.18 (0.0001)
Observations	1,672	1,672	1,672	1,672	1,672	1,672	1,531	1,672	1,672	1,672
Number of agents	63	63	63	63	63	63	57	63	63	63
Adjusted R-squared	0.2794	0.2852	0.2807	0.2992	0.2944	0.3124	0.2277	0.2843	0.2795	0.3272
Individual FE	No	No	No	No	No	No	No	No	No	No

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. All regressions also include a constant, working hours, share in peak hours, calls per FTE, a linear time trend, and fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level. The F-Statistic and corresponding p -value corresponds to a t -test of the main and interaction effects of all added personality traits and cognitive test scores. Due to partial non-response, the sample used in Column 7 is based on 57 agents instead of 63 agents.

Table A.4: Reciprocity and the effect of tenure on the returns to training

	(1)	(2)
Treatment dummy	0.0860*** (0.0133)	0.0685*** (0.0100)
Treatment \times reciprocity	0.0158** (0.0057)	0.0152* (0.0083)
Treatment \times tenure		0.0038 (0.0034)
Constant	-0.9764*** (0.0628)	-0.9763*** (0.0629)
Observations	1,672	1,672
Number of agents	63	63
Adjusted R-squared	0.6127	0.6129
Individual FE	Yes	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. All regressions also include a constant, working hours, share in peak hours, calls per fulltime equivalents (FTE), a linear time trend, and fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level.

Table A.5: Descriptive statistics

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Below median reciprocity		Diff (3)–(2)	All	Above median reciprocity		Diff (3)–(2)
		Treat. group	Control group			Treat. group	Control group	
Gender (1=male)	0.2581 (0.4448)	0.3125 (0.4787)	0.2000 (0.4140)	-0.1125 (-0.6977)	0.3438 (0.4826)	0.4286 (0.5136)	0.2778 (0.4609)	-0.1508 (-0.8736)
Age	37.2165 (12.9927)	36.5168 (11.9346)	37.9628 (14.4212)	1.4460 (0.3049)	35.5114 (9.3614)	33.1854 (8.3138)	37.3205 (9.9514)	4.1351 (1.2508)
Tenure (in years)	4.2295 (3.7824)	4.8714 (3.7375)	3.5449 (3.8364)	-1.3265 (-0.9750)	4.1695 (4.2043)	3.8929 (4.2019)	4.3846 (4.3150)	0.4918 (0.3235)
High education	0.3448 (0.4837)	0.4000 (0.5071)	0.2857 (0.4688)	-0.1143 (-0.6289)	0.2759 (0.4549)	0.3077 (0.4804)	0.2500 (0.4472)	-0.0577 (-0.3343)
Performance	0.3593 (0.0854)	0.3673 (0.0857)	0.3507 (0.0871)	-0.0166 (-0.5354)	0.3664 (0.0832)	0.3674 (0.0576)	0.3657 (0.1005)	-0.0017 (-0.0557)
Number of calls	172.7419 (105.6485)	179.3125 (113.9209)	165.7333 (99.5529)	-13.5792 (-0.3524)	219.3438 (122.1191)	195.7857 (128.6003)	237.6667 (117.1972)	41.8810 (0.9612)
Working hours	15.0000 (8.4499)	15.3750 (9.1497)	14.6000 (7.9355)	-0.7750 (-0.2512)	18.2500 (8.3280)	16.6429 (9.8106)	19.5000 (7.0063)	2.8571 (0.9616)
Share Peak-Hours	0.5243 (0.2095)	0.5015 (0.1993)	0.5486 (0.2241)	0.0471 (0.6186)	0.5411 (0.1797)	0.5810 (0.1555)	0.5101 (0.1951)	-0.0709 (-1.1117)
Absenteeism	0.1935 (0.4016)	0.1875 (0.4031)	0.2000 (0.4140)	0.0125 (0.0852)	0.0312 (0.1768)	0.0000 (0.0000)	0.0556 (0.2357)	0.0556 (0.8787)
Training incidence	0.1935 (0.4016)	0.1250 (0.3416)	0.2667 (0.4577)	0.1417 (0.9809)	0.1875 (0.3966)	0.0714 (0.2673)	0.2778 (0.4609)	0.2063 (1.4886)
Leaver	0.5806 (0.5016)	0.6875 (0.4787)	0.4667 (0.5164)	-0.2208 (-1.2357)	0.4688 (0.5070)	0.4286 (0.5136)	0.5000 (0.5145)	0.0714 (0.3899)
Positive reciprocity	3.6344 (0.3883)	3.5417 (0.4367)	3.7333 (0.3137)	0.1917 (1.3949)	4.7500 (0.3168)	4.7857 (0.3096)	4.7222 (0.3284)	-0.0635 (-0.5561)
Positive reciprocity (std.)	-0.8674 (0.6025)	-1.0113 (0.6776)	-0.7139 (0.4867)	0.2974 (1.3949)	0.8635 (0.4915)	0.9189 (0.4804)	0.8204 (0.5095)	-0.0985 (-0.5561)
Conscientiousness	12.2258 (1.2304)	12.2500 (1.3416)	12.2000 (1.1464)	-0.0500 (-0.1112)	13.0938 (1.4670)	12.9286 (1.3281)	13.2222 (1.5925)	0.2937 (0.5554)
Extraversion	11.8065 (1.7780)	12.6875 (1.0145)	10.8667 (1.9591)	-1.8208** (-3.2804)	12.3750 (1.9134)	12.6429 (2.0232)	12.1667 (1.8550)	-0.4762 (-0.6925)
Agreeableness	12.2903 (1.5747)	11.9375 (1.9822)	12.6667 (0.8997)	0.7292 (1.3034)	13.2188 (1.6798)	13.2857 (1.4899)	13.1667 (1.8550)	-0.1190 (-0.1958)
Openness to experience	10.0645 (1.9653)	10.3125 (2.0238)	9.8000 (1.9346)	-0.5125 (-0.7197)	10.6875 (1.8912)	10.7857 (1.8472)	10.6111 (1.9745)	-0.1746 (-0.2551)
Neuroticism	7.7742 (1.9272)	7.9375 (2.0156)	7.6000 (1.8822)	-0.3375 (-0.4810)	6.8125 (2.6813)	6.5000 (2.7104)	7.0556 (2.7110)	0.5556 (0.5751)
Cognitive test score	0.4425 (0.2683)	0.4583 (0.2950)	0.4231 (0.2417)	-0.0353 (-0.3464)	0.4702 (0.2225)	0.5000 (0.2546)	0.4444 (0.1959)	-0.0556 (-0.6520)
Negative reciprocity	2.4946 (0.6547)	2.5417 (0.7589)	2.4444 (0.5443)	-0.0972 (-0.4074)	2.4688 (1.0772)	2.5714 (0.9734)	2.3889 (1.1730)	-0.1825 (-0.4695)
Locus of Control	8.8387 (1.2675)	8.6875 (1.1955)	9.0000 (1.3628)	0.3125 (0.6798)	8.5938 (1.7754)	9.0714 (2.0178)	8.2222 (1.5168)	-0.8492 (-1.3605)
Loyalty	1.2516 (0.5316)	1.1625 (0.5620)	1.3467 (0.4984)	0.1842 (0.9628)	1.0062 (0.8293)	1.1286 (0.7508)	0.9111 (0.8950)	-0.2175 (-0.7303)
Observations	31	16	15	31	32	14	18	32

Note: Standard deviations are in parentheses in Columns 1 to 3, and 5 to 7; *t*-statistics in Columns 4 and 8.

Table A.6: The contemporaneous effect of peers in training on own performance

	(1)
Share of peers in training	-0.0350 (0.0645)
Calls per FTE	0.0000 (0.0000)
Constant	-1.0869*** (0.0629)
Observations	1,672
Number of agents	63
Adjusted R-squared	0.6022
Individual FE	Yes

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable: $\log(y_{it})$. Share of peers in training is defined as the number of co-workers in an agent's team, divided by the overall number of agents in a team. All regressions also include a constant, working hours, share in peak hours, calls per fulltime equivalents (FTE), a linear time trend, and fixed effects for the week in which agents participated in the survey. Standard errors are clustered at the team level.