When 3+1 > 4:
Gift Structure and Reciprocity in the Field

Duncan Gilchrist[1], Michael Luca[2], and Deepak Malhotra[3]

This version: May 28, 2014

Abstract

Do higher wages elicit reciprocity and hence higher productivity? In a field experiment with 266 employees, we find that paying above-market wages, per se, does not have an effect on productivity relative to paying market wages. However, structuring a portion of the wage as a clear and unexpected gift (by offering a raise with no further conditions after the employee has accepted the contract – with no future employment) does lead to higher productivity for the duration of the job. Targeted gifts are more efficient than hiring more workers. However, the mechanism makes this unlikely to explain persistent above-market wages.

[1] dgilchrist@hbs.edu, Harvard University
[2] mluca@hbs.edu, Harvard Business School
[3] dmalhotra@hbs.edu, Harvard Business School

We thank Kristiana Laugen for exceptional research assistance.

1. Introduction

Economists have long recognized that employees are often paid more than the market clearing wage, and that unemployed workers do not bid wages down to the point where supply equals demand. The neoclassical explanation for this phenomenon comes in the form of efficiency wage theories, generally arguing that employees will work harder when they receive high wages because they do not want to lose a high-paying job (Katz 1986). This type of mechanism relies on repeated interactions between employers and employees. In one-time jobs without any consideration for future employment, the neoclassical model would argue that efficiency wages do not increase productivity.

At the same time, a robust literature based in behavioral economics demonstrates that people care about fairness, and that these fairness considerations may create incentives for reciprocation. If you give a gift to someone, that person might reciprocate – even in a one-shot game with no potential for future interaction (Fehr and Gächter 2000, Falk et al. 2008). This principle has been implemented in field settings as well. For example, Falk (2007) shows that including a small gift in a fundraising letter leads to higher donation levels.

In the context of labor economics, fairness concerns and reciprocity have been offered as an explanation for efficiency wages (Akerlof 1982, Akerlof and Yellen 1990, Fehr et al. 2009). If employees view high wages as a gift, then they may reciprocate by working harder even though there is no financial incentive to do so. The thrust of this argument is that the market wage serves as a reference point, and employees will reward positive deviations from this reference point – even in a one-shot employment contract with no career concerns.

Do employees work harder when they are paid more? Laboratory experiments have mostly shown that paying an unconditional bonus before the work starts causes workers to reciprocate by working harder (e.g. Fehr et al. 1993, Fehr and Gachter 2000, Hannan et al. 2002, Charness 2004). In seminal work in the field, Gneezy and List (2006) look at two field settings, hiring roughly 40 workers to test whether paying higher than market wages increases output in a library data entry task and door-to-door fundraising. Specifically, Gneezy and List compare the output of workers assigned to a "gift" treatment, in which workers are hired at a low wage and then offered a raise immediately before starting work, with the output of workers assigned to a “nongift” treatment, in which workers are hired at and paid the low wage. They find that workers who receive the “gift” (i.e., the additional money) exert higher productivity for the first few hours of the task but that the effect wears off after a few hours and that, in their case, the temporary increase in productivity does not justify its cost.

However, because prior research has not varied the base wage, the set-up cannot identify whether reciprocity is triggered because wages are above the market rate (which is the argument laid out in the literature that uses reciprocity to explain observed above-market wages) or because workers suddenly receive the gift of an unexpected raise after having already agreed to a job. In other words, wage amount (high vs. low) is confounded with wage structure (surprise vs. no surprise). In this paper, we offer a large-scale gift exchange experiment in a field setting where we vary both whether or not a worker receives an unexpected raise before starting work as well as the base wage offered to potential hires. This subtle difference allows us to differentiate between how unexpected “gifts” and above-market wages affect performance. The closest

Our results show that the structure of the gift is central to generating reciprocity: simply hiring at and paying workers a high wage ($4) has no effect at all compared with hiring at and paying workers a low wage ($3). However, hiring workers at a low wage and then offering them an unexpected raise ($3+$1) significantly increases performance. One common model of reciprocity in labor markets (Akerlof 1982, Akerlof and Yellen 1990) assumes that high wages are the determinant of reciprocation – a model that does not differentiate between our 3+1 and 4 offers. Field experiments have similarly not differentiated between these two treatments.

Our experiment takes place on oDesk, an online labor market with several million registered contractors. Using the oDesk platform allows us to vary wages and gifts in a setting where workers are accustomed to tasks like ours – note that ours is a natural field experiment, meaning that the employees do not know that they are part of an experiment. Additionally, using oDesk allows us to hire workers with varying levels of work experience on the platform so that we can analyze precisely how experience and the reference point it sets interact with reactions to the gift. Importantly, the oDesk marketplace allows us to conduct targeted hiring by directly inviting workers to take up our job instead of simply posting a job publicly and waiting for applications. This means we are able to hire workers at different base wages (without individuals knowing how much others have been paid) so that we can test whether it is the base wage or the unexpected gift that affects performance. Normally selection would be a concern when hiring at different wages but the take up rate is 95% among workers with prior experience. Furthermore, recruiting through oDesk means we are additionally able to compare and control for the entire oDesk work histories of our employees.

Our experimental design, which we describe in more detail in the next section, proceeds by hiring three groups of oDesk workers for a data entry task, all of whom request wages less than $3 per hour according to their oDesk profiles. We are clear in our recruitment messages that this is a one-time job. The first group is hired at $3 per hour (i.e., the "3" condition). The second group is also hired at $3 per hour, but before starting work group two is then told that we unexpectedly have extra money in the budget and will pay an extra $1 per hour, so that the total they will receive is $4 per hour ("3+1"). The third group is hired directly at $4 per hour so the fact that we are paying them the higher above-market wage does not signal a "gift" in a salient way ("4"). To ensure the validity of the results, we choose a data entry task (entering CAPTCHAs, to be described in more detail later) that is fairly common on online labor markets and we only recruited workers who self-reported data entry as a specialty on their oDesk profiles.

Consistent with the notion of reciprocity, we find that higher wages that include a surprise gift after hiring the employee (3+1) lead to higher and more persistent productivity across our task relative to the other two groups (3 and 4). More specifically, paying $3+$1 yielded a 20% increase in productivity compared to paying $4, with no additional cost. Compared to paying $3, paying $3+$1 resulted in a 20% increase in productivity with a 33% increase in cost. Notably, the effect is strongest for the workers for whom the gift is most likely to be salient: employees with the lowest prior wages, the most experience (who are more familiar with the standard wage structure on oDesk), and those who have worked most recently. Among this first group (i.e., those with the lowest prior wages), the percentage increase in productivity is actually greater than the percentage increase in cost even when we compare $3+$1 with $3. However, we find that varying the base wage from $3 to $4 in the original contract has no statistically distinguishable effect on productivity – in fact, the point estimate of the effect is 0.

These results help to shed light on the situations in which we should expect to see reciprocity in labor markets. Prior work has focused largely on students with presumably little or no experience with the task at hand, and found small effects. In contrast, we find a large, persistent effect driven by workers with prior experience doing similar work in a natural field setting. Moreover, ours is the first to include the 4 condition – allowing us to better understand the underlying mechanisms driving reciprocity.

Altogether, our results suggest that unexpected gifts (3+1) targeted at those who least expect them can increase productivity – sometimes in a cost-effective manner – but that such reciprocity is due to the unexpected nature of the gift and not simply due to the fact that a worker is receiving a higher, above-market wage. Indeed our results suggest that reciprocity and gift giving are unlikely to explain most efficiency wages as we see them in the world around us because high wages written into the initial wage contract do not seem to elicit any productivity increase.

2. Experimental Design

Our experimental methodology proceeded in three steps. First, we selected a sample of oDesk workers and notified the treated workers that we had a job for them and invited them to accept our job offer. Then, treated workers who accepted our job offers were, if appropriate (i.e., in the 3+1 condition), notified of a change in their wages. Finally, all treated workers were sent a link to a website where they could complete our data entry task. We describe these steps in more detail below.

2.1 Sample

Our sample selection began by restricting the universe of oDesk workers to those who are experts in data entry by requiring that (a) worker profiles are returned by a keyword search for “data entry” and (b) workers classify themselves as Administrative Support workers with a Data Entry specialty.[4] We further restricted the sample to workers that list

[4] We note that oDesk specialties are self-reported so neither of these two restrictions required a worker to actually have any experience in data entry on the oDesk platform (or elsewhere).

a requested hourly wage of between $2.22 and $3.33. Since oDesk charges workers a 10% fee on gross earnings, this restriction amounts to restricting the sample to workers requesting net hourly wages of between $2 and $3. Finally, we restricted the sample to workers that (1) had logged into oDesk within the last 30 days, (2) are independent contractors (unaffiliated with a larger company), and (3) are listed as "oDesk ready", which means they have passed a multiple choice exam that tests their knowledge of the oDesk interface. A total of 17,482 workers satisfied these joint criteria at the time of data collection.

From this set of workers, we randomly selected 540 workers and allocated them randomly across our 3 treatments.[5] Figure 1, Panel A presents the initial recruitment messages with which we invited selected workers to take up our task.

2.2 Treatments

Workers were randomized into three treatment groups – Figure 2 presents our experimental design. Workers in two of the treatment groups were offered $3 per hour while workers in the third group were offered $4 per hour. (Technically, we would pay $3.33 and $4.44, respectively, because 10% of gross wages go to oDesk.)

We refer to $4 per hour as the gift condition, since these employees are receiving more than $3 per hour (despite being pulled from the same set of potential employees). This treatment largely goes ignored in the literature, despite the fact that the literature often describes high wages as a gift. We feel this treatment is important as it allows us to tease out mechanisms behind the gift effect.

The two groups who had been offered $3 per hour are split into salient gift (3+1) and no gift (3). After accepting our job offer, workers in all treatments were reminded of the task instructions and presented with a link to a webpage where the task was located. Workers in the salient gift treatment (3+1) were additionally notified in the same message that we "have a bigger budget than expected … [and that] we will pay … $4 per hour instead of $3 per hour"[6]. Figure 1, Panel B presents the messages we sent to workers to

[5] We oversampled the $3/hour treatment because those workers participated in a trust-building exercise independent of this experiment that took place after this experiment was completed.

[6] This message was phrased so as not to emphasize any sacrifice the firm was making in paying a higher wage. However, our design does not allow us to analyze the impact of the specific wording of this message.

let them know we had agreed to hire them, and Figure 2 summarizes the experimental set up.

As we mention above, our task asked workers to correctly enter as many CAPTCHAs as they can in the four hours allotted. Figure 3 presents a screenshot of the task itself, as seen by workers. CAPTCHA is an acronym for "Completely Automated Public Turing test to tell Computers and Humans Apart", which is a system that asks you to transcribe a word or phrase that is presented to you as a picture. Many online companies use CAPTCHAs to prevent automated software from easily accessing information or making decisions without a human being involved. For example, Ticketmaster requires potential purchasers to enter a CAPTCHA before purchasing tickets in order to stop a person from writing programs to repeatedly buy tickets (which is something that a scalper may otherwise do). On online labor markets such as oDesk, there is a high level of demand for people to do data entry (and in fact, even specifically to enter CAPTCHAs), which means that our task would come across as a reasonably natural request.

In our analysis, we exclude results from 12 workers who did not complete the 4 hours of work, as well as 24 workers in the initial wave who were able to complete more than 4 hours of work due to a technical glitch that allowed them to exceed the time limit.[7] We are left with 230 employees, all of whom are included in our analysis. Our take up rate was 95% among workers with prior experience and 22% among inexperienced workers.

In Table 1, we present worker characteristics before the experiment to verify that randomization was successful and ensure that there is no apparent selection among takers. We first present statistics for all workers, in Panel A, and then statistics on the subsample of experienced workers (workers with at least one prior job), in Panel B. Within each panel, the first set of three columns present statistics, separated by treatment group, for those who did not accept our job offer, and the second set of three columns presents analogous statistics for those who did take up our job. The statistics presented

[7] There was no difference across the three treatment groups on the likelihood of employees working for more or less than 4 hours; we excluded these workers because allowing employees to work for different lengths of time makes it harder to compare total productivity across employees. Including these employees in the analysis does not change our baseline results, however; see Appendix Table 1.

are the mean of a characteristic by subgroup, and then below, in parentheses, the p-value from a T-test comparing a covariate from a given subgroup with the analogous 3+1 subgroup. Although our overall take up rate was 46%, job acceptance rates were similar across treatment groups – as were other worker characteristics. The only notable difference (which is still not statistically significant) across groups is that the number of prior jobs is lower in the 4 group than in the 3+1 and 3 groups. In aggregate, the data suggests that the experimental design is valid.[8]

3. Results

This section documents our main results.

3.1 Main Effect: 3+1>4=3

The gift literature has posited that high wages may elicit reciprocity, which could in turn rationalize above-market wages even in a one-shot labor employment. This suggests that market wages might be a reference point, and that paying more would elicit higher productivity. In this case, our 4 and 3+1 conditions should elicit the same response, which would be higher than the 3 condition. To our knowledge, we are the first to include a 4 treatment, which allows us to shed light on the conditions under which we should expect workers to reciprocate high wages.

Our main finding is that the response to the 4 and 3+1 treatments is very different. At the same time, the estimates for the 3 and 4 treatments are statistically indistinguishable from one another. The gift only matters *if* the wage is structured in a way that the gift component is made salient (e.g., by presenting it separately, when it is unexpected). Figure 4 presents the main effect. Paying 4 elicits the same amount of productivity as paying 3. However, The 3+1 group correctly entered 146 more CAPTCHAs relative to *both* groups over the course of the task, a 20% increase. This result is statistically significant at the 5% level. Table 2 presents the main effect in regression form, where all measurements are relative to the 3+1 treatment. Column (1)

[8] As a robustness check against selective take up, we repeated the analysis in Table 1 in which we compare characteristics across treatments but this time pooling the 3 and 3+1 treatments since these treatments were hired at the same wage. The results are excluded here but none of the differences between the pooled treatments hired at \$3/hour and the treatment hired at \$4/hour were statistically significant at the 10% level.

presents results from a regression of the total number of completed correct CAPTCHAs on a constant and dummies for wages of 3 and 4, respectively. The point estimates of 3 and 4 are nearly identical, and we cannot reject the hypothesis that they are the same. Column (2) of Table 2 presents results in log form. As before, we find the 3+1 treatment outperforms the other two groups, which are statistically indistinguishable from each other. In column (3), we show that the treatment effect measured in number of correct CAPTCHAs entered is not due to a change in the ratio of correct to incorrect entries. The estimates for all treatments are statistically indistinguishable from one another.

Although our sample is balanced on observables and take up rates were similar across all treatments (see Table 1), we conduct an additional robustness check to verify that our result that 3+1>4 is not driven due to selection induced by hiring at different wages. That is, a potential explanation for why 4 performs worse than 3+1 is that negative-sorting led less-able workers to refuse the job invitation for \$3/hour (into the 3+1 treatment) but those same kinds of workers did accept the job invitation for \$4/hour. Such a scenario could lead us to falsely conclude that the 3+1>4 when really selection was driving the result. (Note, however, that such a scenario is counter-intuitive since usually one would expect that positive selection, which would lead to the opposite conclusion: selective take up would lead better workers to be more likely to accept the higher wage job.) To test whether negative selective take up is driving our result, we run a regression where all non-takers are coded as having completed 0 CAPTCHAs and all takers are coded normally. If the differential yield is driving our result and the incentive provided by 4 is actually equivalent to that provided by 3+1, then 4 should perform at least as well as 3+1 in this regression since here we are estimating an average treatment effect conditional on job invitation instead of an average treatment effect conditional on job take up. We report the results in column 2 of Appendix Table 1. The 3+1 treatment still performs better than 4 and the difference remains statistically significant at the 10% level. We conclude that our results are not driven by selection.

Our main result contrasts with Gneezy and List (2006), Kube et al. (2012), and concurrent but independent work by Esteves-Sorenson and Macera (2013), who all find mild to no effects of gifts in the field when hiring undergraduate students for field experiments (although Kube et al. 2012 find that nonfinancial gifts do have a significant effect). In all of these settings, workers presumably had little or no prior experience doing the task they were assigned. In section 3.2, we analyze how experience interacts with impact of the salient gift in our setting. Consistent with previous work, we find no effect among workers without prior experience. However, we do find large effects for employees with prior data entry experience. Among experienced workers, the 3+1 treatment is always efficient relative to the 4 treatment, and sometimes even efficient relative to the 3 (relative to hiring more workers). Our findings relate also to Cohn et al. (forthcoming), who find evidence of reciprocity among employees who viewed their initial wage as unfair.

3.1.1 Does the effect of a gift wear off?

The final column in Table 2 examines worker performance by treatment over the course of the task. Our task database recorded the timestamp of every CAPTCHA entry, so we are able to examine the time series of responses. We gave workers one week to complete the task and workers were not required to complete the 4 hour task in a single (uninterrupted) session.[9] Thus, in order to break the task time into 4 quarters, we first truncate any gaps between CAPTCHA entry that are longer than 10 minutes to just 10 minutes, and then we break the total time between the first and the last CAPTCHA entry (with truncated breaks) into 4 equal blocks of time.[10] The dependent variable in the regression is the number of correct CAPTCHAs entered in each quarter by each worker (so observations are the worker-quarter level and standard errors are clustered by worker). The un-interacted treatment coefficients in the top rows show that that the salient gift workers completed an average of 32 to 33 more correct CAPTCHAs in the first quarter of the task, and this difference is significant at the 5% level for the 3 group and at the 10% level for the 4 group.

In order to examine the effect of the salient gift over the course of the experiment, we present in the bottom panel of Table 2 the sum of the treatment and treatment-quarter

[9] oDesk's platform logged and billed employers for the hours a worker was logged into oDesk's task-specific interface. Employers were given the option of specifying a time limit, which, in our case, was set to 4 hours.

[10] We chose to truncate gaps of longer than 10 minutes because oDesk takes screenshots of its workers to verify their focus and 10 minutes is the approximate time between screenshots, so a break longer than 10 minutes is likely to represent true time away from the task.

estimates. The salient gift treatment outperformed the other treatments in all quarters by 30 to 40 CAPTCHAs. We infer from these results that the salient gift treatment increased productivity across the length of the task. Levitt and List (2012) suggest that lab evidence on reciprocity may not generalize to the field because of the short-term nature of the response (as found in Gneezy and List 2006). While our results do not speak to the longer-run persistence of the effects of a single gift, they do show that the impact of a gift is not necessarily as ephemeral as some of the previous evidence has shown.

3.2 Does experience matter?

The main effect shows that high wages increase productivity only when the wage comes as a surprise. In Table 3, we examine whether the impact of the salient gift differs according to a worker's experience. If the mechanism for the impact of the gift is indeed its saliency then one would expect that the impact of the salient gift would be higher for workers with stronger priors on what to expect; e.g., workers with more experience. Of course, there are other reference points that may be shaping perceptions and productivity as well. We find that the impact of the salient gift is largest for experienced employees and, among experienced employees, those with more prior jobs and those who have worked most recently.

In the first specification (column 1), we regress the number of completed correct CAPTCHAs on dummies for the non-salient-gift treatments, a dummy for whether that worker has prior experience on oDesk, and the experienced dummy interacted with the non-gift treatments. The estimates on the base treatment indicators – i.e., the treatment effects for the inexperienced – are not significantly different from 0. However, the treatment effects for the experienced workers – the addition of the baseline treatment effect and the treatment-experienced interaction term, presented in the bottom panel – are smaller and statistically significantly different from 0, indicating that our estimated salient gift effect is concentrated among experienced users. See Figure 5 for a visual presentation of this result.

Columns (2) and (3) in Table 4 further narrow the sample of users who have characteristics associated with a large reaction to the salient gift. Column (2) repeats the analysis in column (1) but restricts the sample to experienced workers and examines whether the number of prior jobs affected the impact of the treatment by splitting the analysis along workers that have had more prior jobs than the median of prior jobs (which was 6). Although we are unable to reject the hypothesis that the gift effect is the same in both groups (the interaction terms are all insignificant), we see that the salient gift effect has a larger magnitude among the workers with the most experience by comparing the estimates in the top rows with those in the bottom panel. This is shown graphically in Figure 6. Finally, in column (3), we present results from an analogous regression examining how the salient gift effect differs by how recently a worker has worked. We find that the effect is most pronounced among those workers who have worked in the last 30 days. Figure 7 presents this result graphically.

3.3 Do prior wages matter?

Table 4 mirrors the analysis in Table 3, this time examining how the level of an employee's prior wages earned in previous jobs on the platform (which are publicly available) interacts with performance. Like experience, prior wages are likely to serves as a reference point that shapes the perceptions of the employees. Overall, the data suggest that the wage offered relative to this possible reference point does increase productivity, although we cannot differentiate between whether the impact of the salient gift differed due to its size relative to prior wages or due to the fact that workers with different prior wages are simply different kinds of workers. Note that this analysis requires us to restrict the sample to only those workers who had prior wages on their profile – that is, experienced workers whose previous employment included hourly (and not just fixed-price) jobs.

In the first two columns, we see that the number of CAPTCHAs completed by workers in the 3+1 treatment is decreasing in their prior wages, and that the effect is the opposite for workers in the \$3/hour and \$4/hour treatment groups. Column (1) groups these two treatments and column (2) separates them. Note that the coefficient on the interaction term for the 3+1 group is statistically significantly different from the coefficient on the interaction term for the other groups, but that the coefficients on the interaction term for the 3 and 4 groups are statistically indistinguishable. These results show that workers with different prior wages responded the same way to 3 and 4 – i.e., that workers with higher prior wages completed more CAPTCHAS – but that the salient gift (3+1) resulted in more CAPTCHA completions for those with lower prior wages. Overall, the gift seems to have the largest effect for low-wage workers.

In columns (3) and (4), we break out the treatment effect by the median of prior wages, which was $1.67/hour, and we present cumulative treatment effects for above median average prior wages in the bottom panel. These results mirror those in columns (1) and (2) and show that the analysis is not driven by the linear specification. In the most general specification in column (4), we see that the salient gift significantly increased mean productivity in the below-median average prior wage group by nearly 389 and 429 CAPTCHAs compared to the 3 and 4 treatment groups, respectively, while the increase in mean productivity in the above-median average prior wage groups is much smaller, around 100, and not statistically significant. This result is presented graphically in Figure 8. An effect of this magnitude suggests that for the workers who previously earned wages below the median of our sample, the salient gift treatment is actually more efficient than the $3/hour treatment in terms of CAPTCHA completions per dollar expenditure – the $1/hour gift increased average CAPTCHA completion from 723 to 1112 for this group, a 54% increase in productivity, while cost increased only 33%. Of course, we note here that in our results, the salient gift is always efficient (and nearly always statistically significantly so) relative to the $4/hour treatment.

4. Discussion

We find that providing employees with an unexpected pay increase can increase productivity – even when there is no prospect for future employment. However, high wages that actually look like the types of efficiency wages we usually see in the field did not have the same impact on productivity as our salient gift treatment, and in fact had no discernible effect at all.

Our experiment that the context within which field experiments take place matters for their interpretation. A recent literature, e.g., Levitt and List (2007), Falk and Heckman (2009), and Camerer (2011), discusses how the generalizability of lab experiments differs relative to that of field experiments. Our work points to issues of generalizability that occur within field experiments.

From a labor and behavioral economics perspective, our results suggest that gift exchange may be unlikely to explain above-market wages but that a surprise gift can still be impactful. This has managerial implications: our findings suggest that increasing an employee's salary may do less to increase performance than spending that money on the employee through other, perhaps unexpected, channels that make salient the gift component.

References

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

Cohn, Alain, Ernst Fehr, and Lorenz Goette. 2009. “Fair Wages and Effort: Evidence from a Field Experiment.” *Management Science, forthcoming*.

Esteves-Sorenson, Costança, and Rosario Macera. 2013. “Revisiting Gift Exchange: Theoretical Considerations and a Field Test.” *working paper*.

Falk, Armin. 2007. “Gift Exchange in the Field.” *Econometrica* 75 (5): 1501–11.

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2008. “Testing theories of fairness—Intentions matter,” *Games and Economic Behavior* 62 (1): 287-303.

Fehr, Ernst, George Kirchsteiger, and Arno Riedl. 1993. “Does Fairness Prevent Market Clearing? An Experimental Investigation.” *Quarterly Journal of Economics* 108 (2): 437–59.

Fehr, Ernst and Simon Gachter. 2000. “Fairness and Retaliation: The Economics of Reciprocity.” *Journal of Economic* Perspectives 14: 159-181.

Fehr, Ernst, Lorenz Goette, and Christian Zehnder. 2009. “A Behavioral Account of the Labor Market: The Role of Fairness Concerns.” *Annual Review of Economics* 1: 355-384.

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

Hannan, R. Lynn, John Kagel, and Donald Moser. 2002. “Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior.” *Journal of Labor Economics* 20 (2): 923-951.

Herz, Holger and Dmitry Taubinsky. “Market Experience is a Reference Point in Judgments of Fairness,” *working paper.*

Katz, Lawrence. 1986. “A Behavioral Account of the Labor Market: The Role of Fairness Concerns,” *working paper.*

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2012. “The Currency of Reciprocity: Gift Exchange in the Workplace.” *American Economic Review*. 102(4): 1644-62.

Levitt, Stephen D., and John A. List. 2007. “What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?” *The Journal of Economic Perspectives.* 21(2): 153–174

Figure 1: Job offer messages
Panel A: Recruitment messages

All treatments
We are currently looking to hire a group of people to help with simple data entry. The job consists of looking at a photograph of a word, and typing that word into the space provided. This is a four-hour job, with the goal of entering as much data as possible while minimizing the number of mistakes. Specifically, we need as many correctly entered words as possible in four hours because we need the data for a future task and only correct entries can be used. You will have seven days to complete the task.

You will be paid $3 ($4) per hour. Therefore, your total payment for the four hours will be $12 ($16). We hope you will accept this job.

Panel B: Acceptance messages

Treatments: 3 & 4
Great, you are hired.

By the way, we want you to know that this is a one time job; we do not expect to have more work in the future.

Below, you will find a link to a page where you will do the data entry. As we mentioned, the job consists of looking at a photograph of a word, and typing that word into the space provided. Please enter words for four hours, after which you will be ineligible to receive further pay. Finally, please take no more than a week. We will not accept work done more than seven days after you receive this assignment.

Link to job: **here**

Treatment: 3+1
Great, you are hired. As it turns out, we have a bigger budget than expected. Therefore, we will pay you $4 per hour instead of $3 per hour, so the total you can earn is $16.

By the way, we want you to know that this is a one time job; we do not expect to have more work in the future.

Below, you will find a link to a page where you will do the data entry. As we mentioned, the job consists of looking at a photograph of a word, and typing that word into the space provided. Please enter words for four hours, after which you will be ineligible to receive further pay. Finally, please take no more than a week. We will not accept work done more than seven days after you receive this assignment.

Link to job: **here**

Figure 2: Experimental design

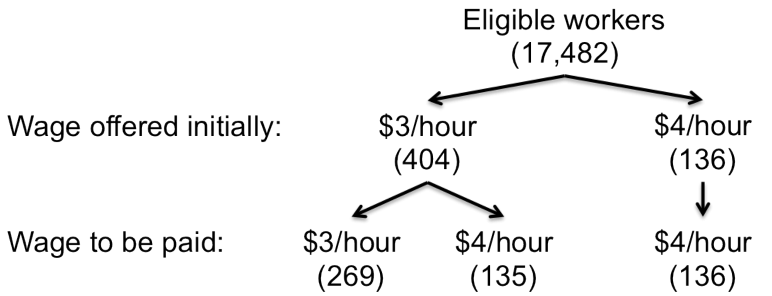


Figure 3: CAPTCHA task

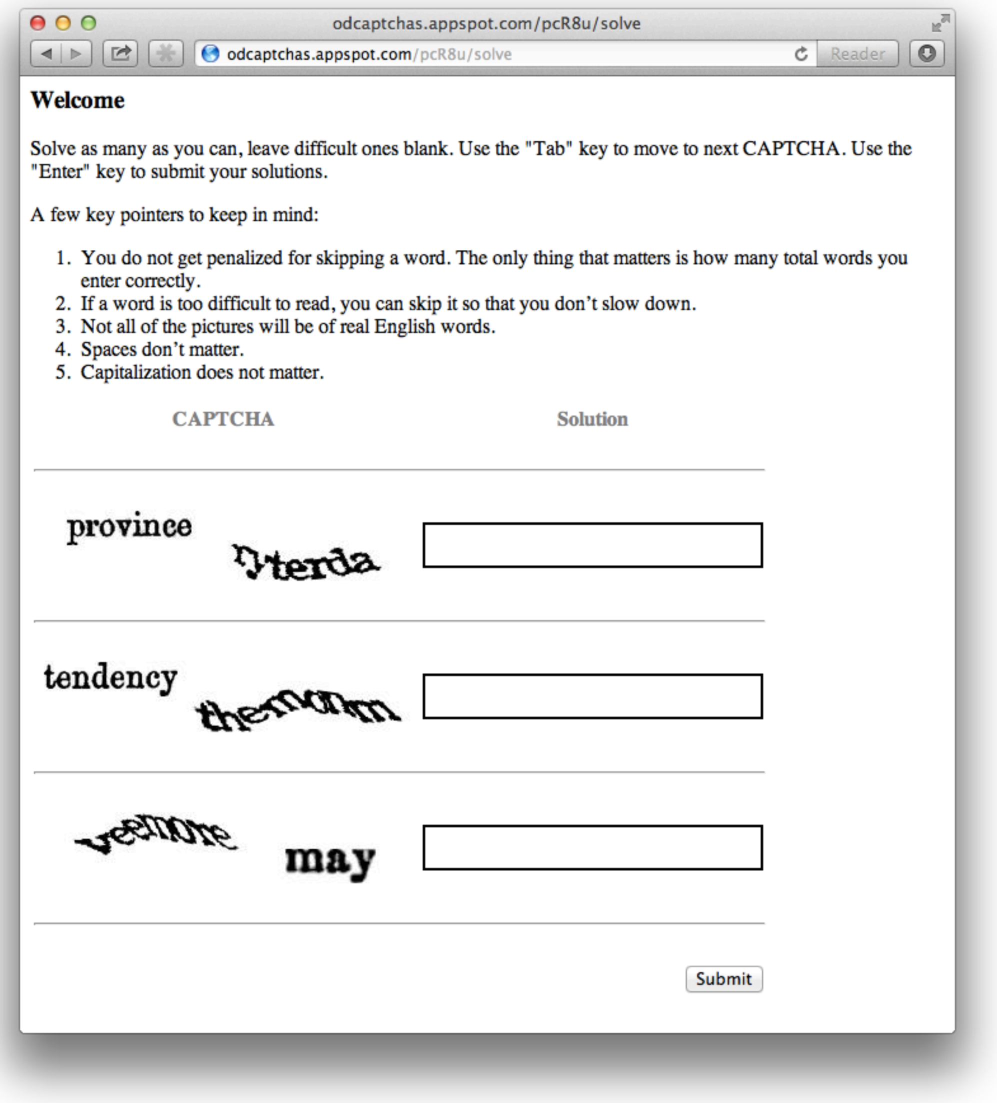


Figure 4: The gift leads to higher productivity than either high or low base wages

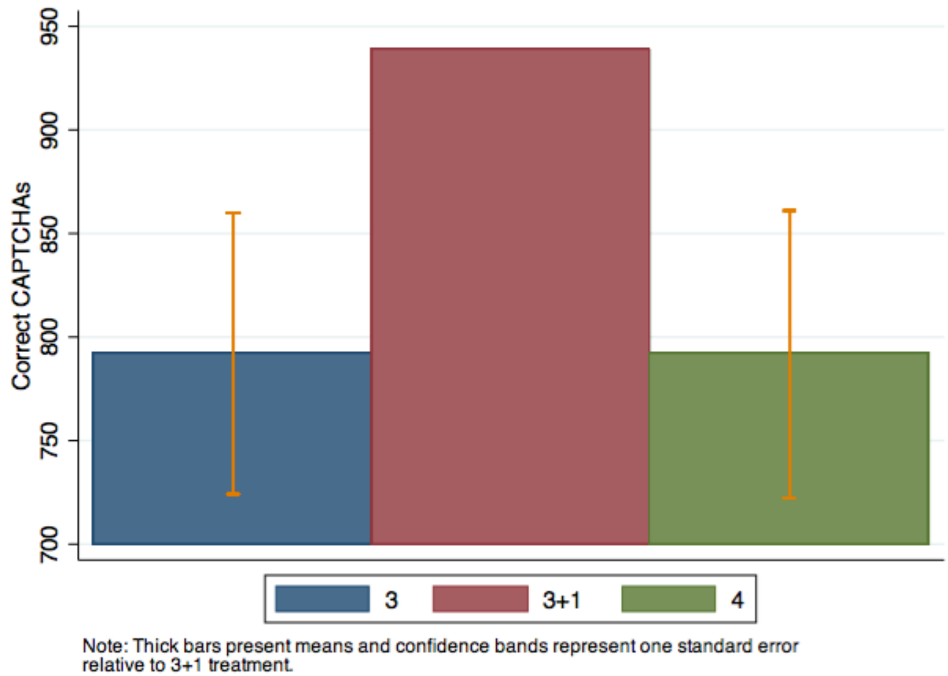


Figure 5: Effects are more pronounced for those with any experience

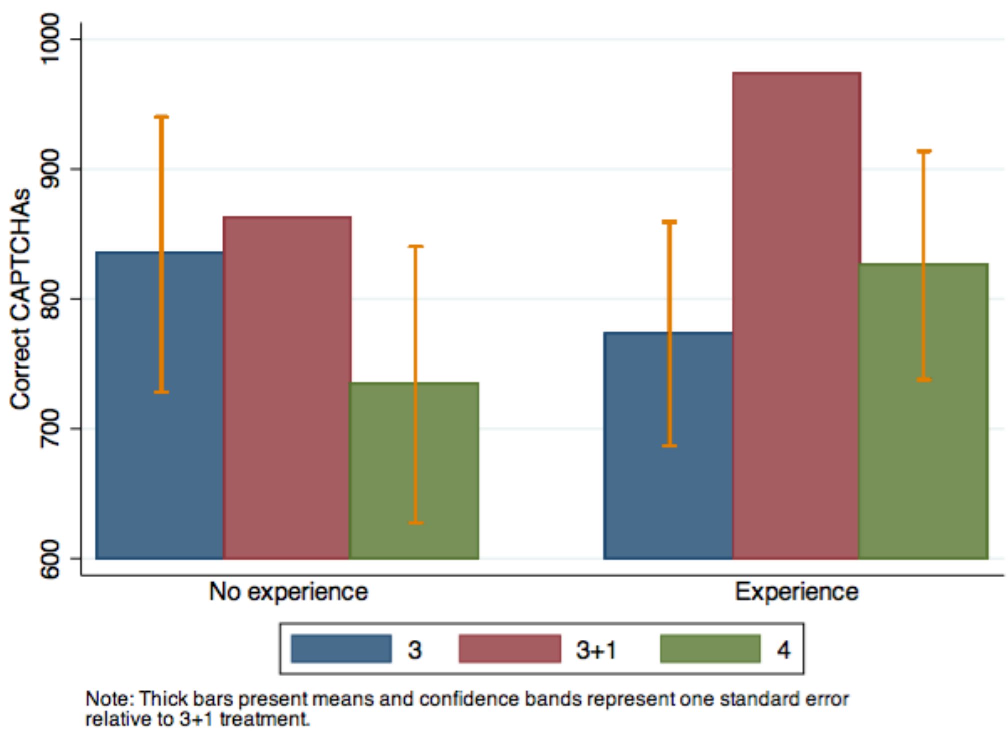


Figure 6: Effects are more pronounced for experienced workers with more prior jobs

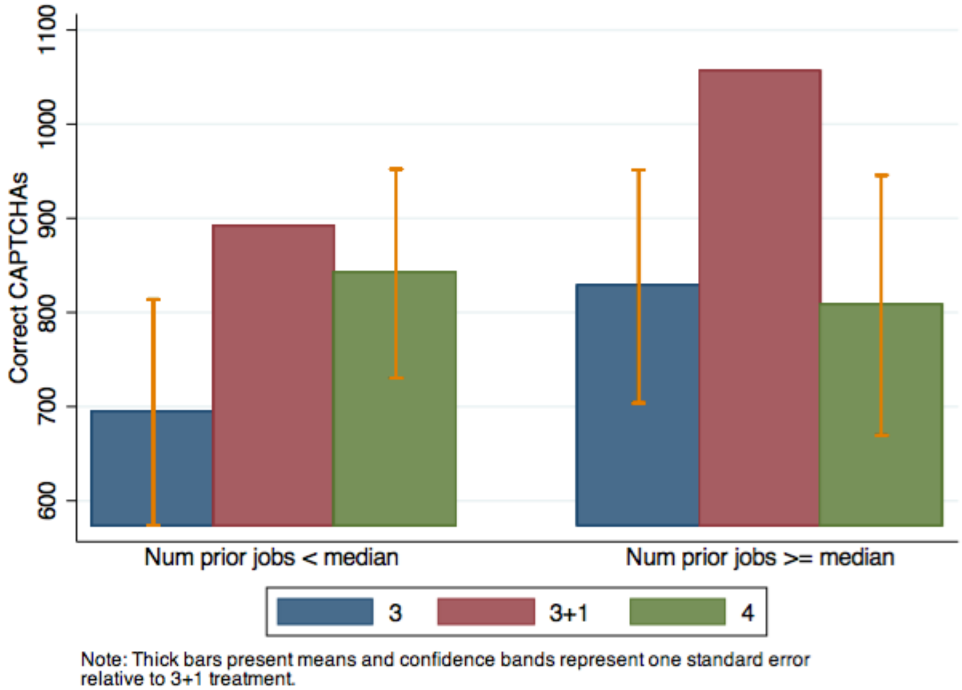
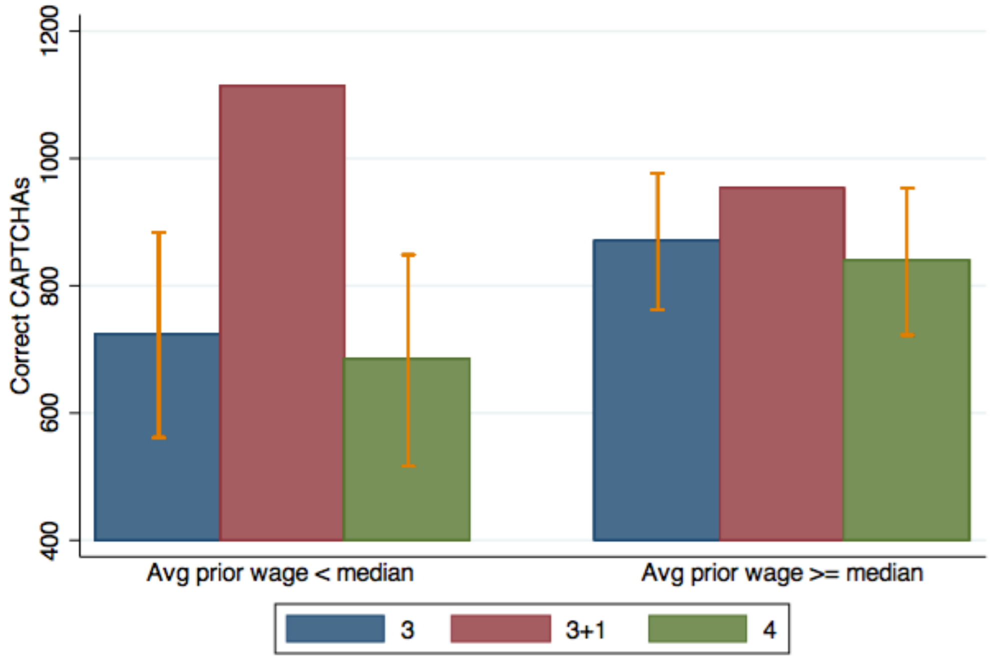


Figure 7: Effects are more pronounced for experienced workers with recent experience

Figure 8: Effects are more pronounced for experienced workers with lower prior wages



Note: Thick bars present means and confidence bands represent one standard error relative to 3+1 treatment.

Table 1: Worker characteristics

Panel A: All workers

Treatment: Wage =	3	3+1	4	3	3+1	4
Taker	No	No	No	Yes	Yes	Yes
Experienced	0.021	0.048	0.029	0.691	0.69	0.629
	(0.301)		(0.527)	(0.987)		(0.482)
Number of prior jobs	0.162	0.238	1	8.082	7.397	4.548
	(0.881)		(0.194)	(0.719)		(0.185)
Wage requested	2.646	2.678	2.666	2.741	2.738	2.815
	(0.675)		(0.884)	(0.972)		(0.392)
N	142	63	69	110	58	62
Take up rate				44%	48%	47%

Panel B: Experienced workers

Treatment: Wage =	3	3+1	4	3	3+1	4
Taker	No	No	No	Yes	Yes	Yes
Number of prior jobs	7.667	5	34.50	11.70	10.72	7.231
	(0.757)		(0.0231)	(0.703)		(0.236)
Wage requested	3.053	2.777	2.775	2.761	2.750	2.882
	(0.585)		(0.998)	(0.909)		(0.234)
Avg prior wage	1.769	1.494	2.233	1.706	1.957	1.908
	(0.716)		(0.397)	(0.191)		(0.828)
Avg prior rating	4.957	4.503	4.615	4.565	4.548	4.495
	(0.285)		(0.802)	(0.913)		(0.759)
Worked in last 30 days	0.333	0	1	0.447	0.525	0.410
	(0.314)		(0.0301)	(0.429)		(0.311)
N	3	3	2	76	40	39
Take up rate				96%	93%	95%

Notes:Variable means are presented in the first row, followed by p-values from a T-test comparing a given sample with the analogous 3+1 treatment in the second row. 26 experienced workers did not have an average prior wage because all their prior jobs were fixed price contracts. These workers are excluded from the calculation of statistics using "Avg prior wage" above: for these statistics there are only respectively 66, 33, and 30 observations in the taker 3, 3+1, and 4 groups Similarly, 8 experienced workers did had have prior ratings because their first and only contract was ongoing at the time of the experiment, so there are only respectively 72, 38, and 37 observations in the taker 3, 3+1, and 4 groups used in calculating those statistics.

Table 2: Performance

	(1) Correct	(2) log(Correct)	(3) Perc. correct	(4) Correct rate
Wage = 3+1	omitted	omitted	omitted	omitted
Wage = 3	-146.4**	-0.286***	-1.224	-32.97**
	(67.93)	(0.100)	(1.433)	(16.23)
Wage = 4	-146.8**	-0.214**	-0.797	-32.44*
	(69.36)	(0.0990)	(1.335)	(16.54)
Quarter = 1				omitted
Quarter = 2				16.12*
				(8.698)
Quarter = 3				18.33*
				(11.02)
Quarter = 4				20.64
				(13.62)
Wage = 3 x Quarter 2				2.161
				(10.09)
Wage = 3 x Quarter 3				-6.182
				(13.02)
Wage = 3 x Quarter 4				-10.52
				(15.05)
Wage = 4 x Quarter 2				-7.685
				(10.16)
Wage = 4 x Quarter 3				-8.070
				(12.76)
Wage = 4 x Quarter 4				-1.880
				(15.42)
Constant	938.5	6.754	85.33	219.9
	(55.04)	(0.0565)	(0.980)	(12.87)
Treatment effects by quarter:				
[Wage = 3] + [Wage = 3 x Quarter 2]				-30.81
				(18.94)
[Wage = 3] + [Wage = 3 x Quarter 3]				-39.15**
				(19.77)
[Wage = 3] + [Wage = 3 x Quarter 4]				-43.49**
				(19.15)
[Wage = 4] + [Wage = 4 x Quarter 2]				-40.13**
				(18.56)
[Wage = 4] + [Wage = 4 x Quarter 3]				-40.51**
				(20.22)
[Wage = 4] + [Wage = 4 x Quarter 4]				-34.32*
				(20.50)
Observations	230	230	230	920
Adjusted R-squared	0.017	0.017	-0.006	0.0014

Notes: Columns 1 through 3 present results from regressions on worker-level observations, while column 4 presents results from regressions conducted on observations at the worker-quarter level. Treatment effects by quarter in the bottom panel are computed by taking the linear combination of a treatment dummy and the interaction of a treatment dummy with a quarter dummy. Standard errors for column 4 are clustered at the respondent level. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Experience

	(1) Correct	(2) Correct	(3) Correct
Wage = 3+1	omitted	omitted	omitted
Wage = 3	-27.34	-197.0	-120.9
	(105.3)	(120.1)	(100.9)
Wage = 4	-127.8	-49.71	-65.93
	(105.5)	(110.9)	(103.5)
Experienced	111.2		
	(104.3)		
Wage = 3 x Experienced	-172.5		
	(136.2)		
Wage = 4 x Experienced	-19.46		
	(137.5)		
Num prior jobs ≥ median		164.1	
		(142.9)	
Wage = 3 x (Num prior jobs ≥ median)		-30.32	
		(172.7)	
Wage = 4 x (Num prior jobs ≥ median)		-197.7	
		(177.1)	
Worked last 30 days			84.66
			(141.5)
Wage = 3 x Worked last 30 days			-161.8
			(170.4)
Wage = 4 x Worked last 30 days			-174.6
			(172.9)
Constant	861.8	853.7	928.6
	(75.23)	(87.19)	(75.58)
Treatment effects by worker type:			
[Wage = 3] + [Wage = 3 x Experienced]	-199.9**		
	(86.46)		
[Wage = 4] + [Wage = 4 x Experienced]	-147.3*		
	(88.27)		
[Wage = 3] + [Wage = 3 x (Num prior jobs ≥ median)]		-227.4*	
		(124.0)	
[Wage = 4] + [Wage = 4 x (Num prior jobs ≥ median)]		-247.4*	
		(138.1)	
[Wage = 3] + [Wage = 3 x Worked in last 30 days]			-282.7**
			(137.3)
[Wage = 4] + [Wage = 4 x Worked in last 30 days]			-240.6*
			(138.5)
Sample	all	experienced	experienced
Observations	230	155	155
Adjusted R-squared	0.035	0.071	0.051

Notes: Treatment effects by worker type in the bottom panel are computed by taking the linear combination of a treatment dummy and the interaction of a treatment dummy with a worker characteristic. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Gift relative to prior wages

	(1) Correct	(2) Correct	(3) Correct	(4) Correct
Wage = 3+1	omitted	omitted	omitted	omitted
Wage = 3	-614.1***	-627.1***	-390.0**	-388.7**
	(206.9)	(214.4)	(157.9)	(161.0)
Wage = 4	-650.1***	-614.8***	-425.9***	-429.2**
	(212.8)	(233.9)	(161.7)	(166.2)
Wage = 3+1 x Avg prior wage	-129.4**	-129.4**		
	(64.23)	(64.49)		
Wage not = 3+1 x Avg prior wage	72.11*			
	(41.47)			
Wage = 3 x Avg prior wage		79.71		
		(51.87)		
Wage = 4 x Avg prior wage		53.60		
		(64.80)		
Avg prior wage ≥ median			-159.3	-159.3
			(165.6)	(166.2)
Wage not = 3+1 x (Avg prior wage ≥ median)			308.6*	
			(182.5)	
Wage = 3 x (Avg prior wage ≥ median)				305.9
				(193.4)
Wage = 4 x (Avg prior wage ≥ median)				314.5
				(202.3)
Constant	1,283	1,283	1,112	1,112
	(187.8)	(188.6)	(146.1)	(146.7)
Treatment effects by avg prior wage:				
[Wage = 3] + [Wage not = 3+1 x (Avg prior wage ≥ median)]			-81.4	
			(101.3)	
[Wage = 4] + [Wage not = 3+1 x (Wage requested ≥ median)]			-117.3	
			(104.8)	
[Wage = 3] + [Wage = 3 x (Avg prior wage ≥ median)]				-82.8
				(107.1)
[Wage = 4] + [Wage = 4 x (Avg prior wage ≥ median)]				-114.7
				(115.3)
Sample	experienced with prior wages			
Observations	129	129	129	129
Adjusted R-squared	0.079	0.072	0.071	0.063

Notes: All columns exclude all inexperienced workers since they do not have average prior wages, as well as 26 experienced workers whose only prior jobs were fixed price contracts. Treatment effects by wage in the bottom panel are computed by taking the linear combination of a treatment dummy and the interaction of a treatment group dummy with (Avg prior wage ≥ median). Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 1: Robustness

	(1) Correct	(2) Correct
Wage = 3+1	omitted	omitted
Wage = 3	-161.6**	-133.4**
	(64.20)	(56.98)
Wage = 4	-156.9**	-115.2*
	(66.22)	(62.43)
Constant	932.9	497.5
	(52.05)	(48.78)
Sample	Taker + Excluded	Taker + Excluded + non-Taker
Observations	266	540
Adjusted R-squared	0.022	0.009

Notes: Column 1 repeats the baseline analysis presented in Table 2, column 1 but includes results for the 36 workers who did not complete exactly 4 hours of work due to a technical glitch. Column 2 includes results for the 36 excluded workers as well as for the 274 workers who did not accept our job offer. Workers who did not take up the job are coded as having completed 0 CAPTCHAs. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.