

Online Appendix

for

No gifts returned: Surprise bonuses reduce productivity in a natural field experiment

Francesco Bogliacino^a, Gianluca Grimalda^b, David Pipke^c

^aUniversità Cattolica del Sacro Cuore, francesco.bogliacino@unicatt.it

^bUniversity of Passau, South East Technological Institute, Institute for Globally Distributed Open Research and Education (IGDORE), gianluca.grimalda@uni-passau.de

^cChristian-Albrechts-Universität zu Kiel, info@dpipke.de

A Supplementary Online Material

In Section A.1 of the Online Appendix, we present our theoretical model. Section A.2 provides additional analyses and figures. Section A.3 contains the experiment instructions, the protocol, and the final questionnaire.

A.1 Theoretical Model: Analyses and Derivations

A.1.1 The model

Drawing upon foundational theoretical contributions (Akerlof, 1982; Rabin, 1993; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), we propose a worker utility function that is composed of four components, echoing the approach of DellaVigna et al. (2022):

$$U_i = w_i - c(e_i) - \zeta \cdot p_d(e_i, B) + K_{f_i} \cdot e_i \quad (2)$$

The first three components in (2) are self-regarding, as they solely depend on variables pertaining to the agent. w_i is the wage earned by the worker. $c(e_i)$ is the cost associated with exerting effort. As standard in the literature, we assume a convex function in which costs increase more than proportionally as effort increases, i.e., $c(0) = 0$, $c'(e) > 0$, $c''(e) > 0$ and $\lim_{e \rightarrow \pm\infty} c(e) = \infty$. The third term captures the disutility derived from the fear of early dismissal, caused by effort falling short of certain standards. It is the product of the probability of dismissal $p_d(e_i, B)$ and of a loss

term ζ .¹³ We assume that p_d has a symmetric shape to $c(e_i)$. When effort is zero, the probability of dismissal equals $\bar{P} \leq 1$ and approaches one in the negative effort domain $\lim_{e \rightarrow -\infty} p_d(e_i, B) = 1$. Higher effort leads to lower probability of dismissal, but this is so at decreasing rates. Formally: $\frac{\partial p_d}{\partial e} < 0$; $\frac{\partial^2 p_d}{\partial e^2} < 0$; $\lim_{e \rightarrow \infty} p_d = 0$, $p_d(0) = \bar{P} \leq 1$.¹⁴ Crucially, we also posit that receiving a bonus reduces p_d , with the worker perceiving the bonus as an indication of the employer’s satisfaction with their effort, hence: $\frac{\partial p_d}{\partial B} < 0$. Additionally, we suggest that the effect of increasing effort e or bonus B on reducing p_d tends to diminish as the value of the other variable increases, implying: $\frac{\partial^2 p_d}{\partial e_i \partial B} \geq 0$.¹⁵ In the follow-up experiment, our specific objective is to gain insight into the processes in workers’ minds upon receiving a bonus.

The fourth component of (2) is other-regarding as it is grounded on both the willingness to affect the entrepreneur’s payoff and on concerns for the coworker’s payoffs. We call it the *social preferences* component.

$$K_{f_i} = \kappa_i + \varphi_i(B) \cdot (w_i - w_{norm}) - \left\{ (I_{w_i > w_j} \cdot \beta_i - I_{w_j > w_i} \cdot \alpha_i) \cdot (w_i - w_j) \right\} \quad (3)$$

In the spirit of Rabin’s (1993) reciprocity model, K_{f_i} is the worker’s perception of the entrepreneur’s “kindness”. The key idea is that an individual is willing to reciprocate a kind action with kindness and a spiteful action with spite (Fehr and Gächter, 1998; Charness, 2004; Abbink and Sadrieh, 2009). K_{f_i} consists of three parts, akin to DellaVigna et al. (2022). The first is a constant κ_i , which may capture unconditional value attributed to the firm’s payoff in reduced form. In addition, we assume that workers’ concerns about the employer’s payoffs are affected by two terms. The first term captures the “gift exchange” Akerlof (1982) component. It expresses the worker’s sensitivity to the firm paying wages above or below an exogenous norm. Such a social norm may coincide with the market equilibrium or the level set in previous bargaining or past interactions. Wages equal to the norm are perceived as neither kind nor unkind (Akerlof, 1982).

The parameter φ_i quantifies the strength of the reciprocity component in the utility function. We assume $\varphi_i \geq 0$. That is, workers generally interpret higher wages above w_{norm} as a sign of greater kindness by the entrepreneur and *vice versa*. Crucially, we assume that φ_i depends on B . We posit that the payment of B may be construed by the worker as a signal that the entrepreneur

¹³We implicitly assume that $p_d(e_i, B)$ is solely dependent on the individual’s receipt of a bonus, and does not factor in whether a coworker receives a bonus, maintaining simplicity in our notation.

¹⁴Job loss arguably represents the most severe fear a worker might face in their employment relationship. However, workers may also harbor lesser fears, such as being criticized, reprimanded, or receiving unfavorable future job references. As our study did not examine beliefs about these milder fears, they are not further addressed here. Notably, a positive signal from an employer, like issuing a bonus, can reduce a worker’s perceived risk of negative consequences for low effort. This situation illustrates a paradox where a bonus, while intended to increase effort through worker reciprocity, might inadvertently signal employer satisfaction with current effort levels. As a result, workers might then lower their perceived risk of adverse outcomes, ranging from criticism to potential job loss.

¹⁵It might be hypothesized that an employee’s perceived risk of dismissal could rise when a bonus is awarded to a colleague but not to them. However, we do not incorporate this hypothesis into our model. This exclusion is due to both analytical simplicity and the empirical challenges in testing this assumption.

is satisfied with the worker’s effort. Such a belief over the employer’s contentment may then lead to a *lower* need to reciprocate the employer’s kindness, thus decreasing φ_i . Formally, $\frac{\partial \varphi_i}{\partial B} \leq 0$.

The third term in K_{f_i} refers to the horizontal inequality between coworkers’ wages. Fehr and Schmidt (1999) posit that individuals dislike both advantageous ($w_i > w_j$) and disadvantageous ($w_i < w_j$) inequality, experiencing *compassion* when earning more than the other worker, and *envy* when earning less. Advantageous and disadvantageous inequalities are weighted by parameters $-\beta_i$ and $-\alpha_i$, respectively. $I_{()}$ is an indicator function that takes the value of 1 if the subscripted condition is fulfilled. Since individuals are normally more sensitive to disadvantageous than advantageous inequality, parameters should satisfy the condition: $\alpha_i \geq \beta_i \geq 0$. Accordingly, individuals perceive the firm as *less* kind if the wage differential is positive.

Another possibility is that individuals are not driven by compassion when earning more than another individual but rather enjoy having higher status than the counterpart (Frank, 1985), or are driven by spite against others. In either case, individuals perceive the entrepreneur as kinder if they reward the agent while not rewarding the counterpart. This can be accommodated into the model by allowing β_i to be negative. Therefore, our model can capture aversion to disadvantageous inequality ($\alpha_i > 0$), compassion ($\beta_i > 0$), status-seeking behavior or spite ($\beta_i < 0$), or indifference to either ($\alpha_i = 0; \beta_i = 0$). These factors, including inequality aversion, do not directly influence the utility function as in Fehr and Schmidt’s (1999) model but indirectly affect the perceived kindness of the employer.¹⁶

In the context of reciprocity theory, the fundamental assumption is that individuals reciprocate in kind – both positive gestures like kindness and negative ones like spite. When $K_{f_i} > 0$, the entrepreneur is perceived as kind, enhancing the individual’s utility in direct proportion to the effort e_i exerted. When K_{f_i} is lower, there is a commensurate reduction in effort. In the extreme case where $K_{f_i} < 0$, the entrepreneur is viewed as unequivocally unkind. Here, the individual’s utility increases as e_i becomes more negative, potentially manifesting as behaviors harmful to the firm’s objectives, such as sabotage.

Given that Equation 3 and e_i enter Equation 2 multiplicatively, the worker will maximize their utility by exerting positive effort as long as their social preferences or sentiments for the firm are positive, i.e., $K_{f_i} > 0$. The utility function for the second worker j is assumed to be identical to that of worker i with reversed indices i and j .

In our experimental setup, wages can only take two values. Either the wage is equal to the fixed wage w , or the worker additionally receives a lump-sum bonus B . Then, the wage of worker i (j) can be expressed as $w_i = w + I_{B_i} \cdot B$ ($w_j = w + I_{B_j} \cdot B$). The utility function for the second worker j is assumed to be identical to that of worker i with reversed indices i and j . Worker i ’s maximization problem can be written as:

¹⁶Our modeling choice is justified by the fact that workers cannot redistribute income directly as they might in the scenarios Fehr and Schmidt (1999) examined. Instead, in our context, workers can express their aversion to inequality solely by reducing their effort, thereby “punishing” the agent (the employer) responsible for this disparity.

$$\operatorname{argmax}_{e_i \in \mathbb{R}} w_i - c(e_i) - \zeta \cdot p_d(e_i, B) + K_{f_i} \cdot e_i \quad (4)$$

The model assumes that a worker’s optimal effort level hinges on balancing the marginal cost of effort and the marginal disutility from fear of dismissal against the marginal effort benefit due to social preferences towards the firm (equation 5).

$$e_i^* = c'^{-1}(K_{f_i} - \zeta \cdot \frac{\partial p_d}{\partial e_i}) \quad (5)$$

Hence, the higher the fear of dismissal and the higher the reciprocity motives, the higher the effort. In Online Appendix Section A.1.3, we derive a condition under which optimal effort e_i^* either increases or decreases upon the payment of the bonus. Given that a bonus will reduce fear of dismissal, thus tending to reduce effort, a sufficient condition is that the overall need to reciprocate the employer’s kindness decreases because of the perception that the employer is satisfied with the worker’s performance. A first factor in the reciprocity component of the utility function models the worker’s perception that the employer is kinder because of the bonus payment – this is captured by the increase in w_i in equation (3). The second factor models the possibility that the overall need to reciprocate the employer’s kindness decreases because of the perception that the employer is satisfied with the worker’s performance. There are various reasons why workers may reduce their need to reciprocate the employer’s perceived kindness. Suppose the reciprocity motive is based on an imperfectly delineated social norm. In that case, the worker might assume that the bonus signals that the employer’s expectation of the appropriate amount of reciprocated effort is lower than what was provided in the morning session. Suppose the reciprocity motive is based on pro-social preferences. In that case, the worker may conceive that the bonus payment signals a lower effort is necessary to reciprocate the employer’s kindness. In either case, a worker will be prompted to reduce their effort. This set of motivations is captured by the assumed non-positive sign of $\frac{\partial \varphi_i}{\partial B}$.¹⁷ If $-\frac{\partial \varphi_i}{\partial B} > B$, we will observe a reduction in effort in the reciprocity component of the utility function, which goes hand-in-hand with the hypothesized reduction in effort due to lower fear of dismissal.

In keeping with standard practice, our model does not incorporate factors such as learning, fatigue, or other time-variant elements that could influence effort levels. Nevertheless, our difference-in-differences analysis of the experimental data, elaborated in Sections 3 and 4, addresses these effects by controlling for unobserved heterogeneity and temporal variations.

A.1.2 Propositions

- Proposition 1: Optimal effort choice

¹⁷One could argue that consistent with the reciprocity hypothesis, issuing a bonus might enhance the strength of the reciprocity component, namely $d\varphi_i/dB > 0$. However, this effect would overlap with the already assumed addition of B to w_i in the $(w_i - w_{norm})$ term of equation (3); hence, we abstract from this.

- The optimal effort level e_i^* is determined by the first-order condition (FOC), where sum of the marginal cost of effort and the marginal disutility linked to the fear of dismissal equals the the marginal effort benefit due to the social preferences towards the firm term. Formally, the optimal level of effort is represented as $e_i^* = c'^{-1}(K_{f_i} - \zeta \cdot \frac{\partial p_d}{\partial e_i})$.
- Proposition 2: Reaction of effort to a bonus
 - The response of optimal effort e_i^* to a change in the bonus B in the absence of wage inequality is governed by differentiating the FOC with respect to B . This response is a function of the sensitivity of perceived fear of dismissal to the bonus, the potential utility loss ζ , and the degree of reciprocity motives φ_i . The formal expression is: $\frac{\delta e_i^*}{\delta B} = \frac{-\zeta \cdot \frac{\partial^2 p_d}{\partial e_i^* \partial B} + \frac{\delta \varphi_i}{\delta B} \cdot (w_i - w_{norm}) + \varphi_i}{-SOC} \cdot I_{B_i}$, where SOC represents the second-order condition of the worker's maximization problem. Hence, the effort response is adversely affected by the potential utility loss due to the fear of negative repercussions ζ and the cross-derivative $\frac{\partial^2 p_d}{\partial e_i^* \partial B}$. Furthermore, the assumed negative sensitivity of φ_i w.r.t. to the bonus B , one mechanism for the contentment effect, multiplied with the difference of the wage $w_i = w + I_{B_i} \cdot B$ can amplify the potential negative impact of a bonus on the level of effort exerted.
 - In summary, the intricate relationship between the sensitivity to perceived risk, potential utility loss, and a diminished inclination for reciprocity due to the bonus suggests that a bonus payment might result in reduced effort. This occurs when the cumulative impact of these factors outweighs the reciprocal motivations embedded in φ_i .
- Proposition 3: Predictions in the presence of wage inequality
 - In scenarios characterized by wage inequality among workers, the model predicts distinct differential responses to bonus payments relative to the relevant baseline, contingent on whether a worker is a recipient or nonrecipient of the bonus. For scenarios where both workers receive a bonus (Double condition), the effort differential, relative to bonus recipients within single-bonus conditions, is expressed as $(\frac{\delta e_i^*}{\delta B} |_{Double} - \frac{\delta e_i^*}{\delta B} |_{ADV}) \cdot dB = \frac{\beta_{T_i}}{-SOC} \cdot dB$, where the subscript ADV indicates a bonus recipient in a single-bonus condition. The differential is positive for $\beta_{T_i} > 0$, i.e., a compassionate response to advantageous inequality.
 - Conversely, in a condition where no worker receives a bonus (Control condition), the effort differential compared to a nonrecipient in a single-bonus condition is $-\frac{\delta e_i^*}{\delta B} |_{DIS} \cdot dB = \frac{\alpha_{T_i}}{-SOC}$, where DIS indicates a nonrecipient in a single-bonus condition. The differential is positive for $\alpha_{T_i} > 0$, i.e., an aversion to disadvantageous inequality, reflecting a lower evaluation of the employer's kindness in such scenarios.

A.1.3 Detailed derivation of the model's propositions

To determine the optimal level of effort e_i^* for worker i , we derive the first-order condition (FOC) from the utility function in Equation 6, leading to the result from **Proposition 1**:

$$\frac{\delta U_i}{\delta e_i} = -c'(e_i) - \zeta \cdot \frac{\partial p_d}{\partial e_i} + K_{f_i} = 0 \iff e_i^* = c'^{-1}(K_{f_i} - \zeta \cdot \frac{\partial p_d}{\partial e_i}) \quad (6)$$

The second-order condition (SOC) for this maximization problem is written in Equation 7.

$$\frac{\delta^2 U_i}{\delta e_i^2} = -c''(e_i) - \zeta \cdot \frac{\partial^2 p_d}{\partial e_i^2} < 0 \quad (7)$$

To analyze the impact of bonus payments on optimal effort, we differentiate the FOC at $e_i^*(B)$

$$\begin{aligned} \frac{\delta^2 U_i}{\delta e_i^* \delta B} = & -\frac{\partial^2 c}{\partial e_i^*} \cdot \frac{\partial e_i^*}{\partial B} - \zeta \cdot \frac{\partial^2 p_d}{\partial e_i^{*2}} \cdot \frac{\partial e_i^*}{\partial B} - \zeta \cdot \frac{\partial^2 p_d}{\partial e_i^* \partial B} + \frac{\delta \varphi_i}{\delta B} \cdot (w_i - w_{norm}) \\ & + \varphi_i \cdot I_{B_i} - I_{w_i > w_j} \cdot \beta_i) \cdot I_{B_i} - I_{w_j > w_i} \cdot \alpha_i I_{B_j} = 0 \end{aligned} \quad (8)$$

Rearranging Equation 8 provides the response of optimal effort to a bonus as written in Equation 9 to prove **Proposition 2**:

$$\frac{\delta e_i^*}{\delta B} = \frac{-\zeta \cdot \frac{\partial^2 p_d}{\partial e_i^* \partial B} + \frac{\delta \varphi_i}{\delta B} \cdot (w_i - w_{norm}) + \varphi_i \cdot I_{B_i} - I_{w_i > w_j} \cdot \beta_i - I_{w_j > w_i} \cdot \alpha_i}{\frac{\partial^2 c}{\partial e_i^*} + \zeta \cdot \frac{\partial^2 p_d}{\partial e_i^{*2}} \cdot I_{B_i}} \quad (9)$$

Equation 9 can also be expressed as in Equation 10.

$$\frac{\delta e_i^*}{\delta B} = \frac{-\zeta \cdot \frac{\partial^2 p_d}{\partial e_i^* \partial B} + \frac{\delta \varphi_i}{\delta B} \cdot (w_i - w_{norm}) + \varphi_i \cdot I_{B_i} - I_{w_i > w_j} \cdot \beta_i - I_{w_j > w_i} \cdot \alpha_i}{-SOC} \quad (10)$$

In the absence of wage inequality, the Equation 10 simplifies to Equation 11.

$$\frac{\delta e_i^*}{\delta B} = \frac{-\zeta \cdot \frac{\partial^2 p_d}{\partial e_i^* \partial B} + \frac{\delta \varphi_i}{\delta B} \cdot (w_i - w_{norm}) + \varphi_i \cdot I_{B_i}}{-SOC} \quad (11)$$

The model indicates that effort response is adversely affected by the potential utility loss ζ and the size of the cross-derivative $\frac{\partial^2 p_d}{\partial e_i^* \partial B}$. Furthermore, the assumed negative sensitivity of φ_i w.r.t. to the bonus B , one mechanism for the contentment effect, multiplied with the difference of the wage $w_i = w + I_{B_i} \cdot B$ can amplify the potential negative impact of a bonus on the level of effort exerted. The interplay between the sensitivity of perceived risk, potential utility loss, and a weaker

perceived need for reciprocal behavior in response to the bonus suggests that a bonus payment can lead to decreased effort if these effects supersede the reciprocity motives encapsulated in φ_i .

Regarding the differential effort responses to bonuses, our model delineates distinct outcomes for scenarios where either one or both workers receive a bonus. The effort disparity between scenarios where both workers (none) receive a bonus, and a situation where only one worker receives a bonus (does not receive a bonus while their coworker does), is captured by Equation 12 (13). These results from **Proposition 3** follow directly from subtracting Equation 10 adapted to different conditions from each other. Here, the subscript T denotes the various justifications (Arbitrary, Productivity, and Needs, as per our experiment) provided for assigning the bonus in single-bonus conditions. Equation 12 elucidates that when both workers receive a bonus, the resultant effort difference, compared to a scenario where only one worker is the recipient in a single-bonus condition, is positive if $\beta_{T_i} > 0$, indicating a presence of compassion. Conversely, in the Control condition, where no worker receives a bonus, the effort difference is greater than that of a non-recipient in a single-bonus condition, provided $\alpha_{T_i} > 0$, i.e., an aversion to disadvantageous inequality and a lower evaluation of the employer's kindness in such scenarios. The variation in effort response to a bonus within the groups of bonus recipients and non-recipients across different treatments, each with its unique justification for bonus allocation, hinges on the relative intensities of aversion to both disadvantageous and advantageous inequality within these treatments, as quantified by α_{T_i} and β_{T_i} , respectively.

$$\left(\frac{\delta e_i^*}{\delta B} \Big|_{Double} - \frac{\delta e_i^*}{\delta B} \Big|_{ADV} \right) \cdot dB = \frac{\beta_{T_i}}{-SOC} \quad (12)$$

$$\frac{\delta e_i^*}{\delta B} \Big|_{Control} \cdot 0 - \frac{\delta e_i^*}{\delta B} \Big|_{DIS} \cdot dB = \frac{\alpha_{T_i}}{-SOC} \quad (13)$$

A.2 Supplementary Analyses and Statistics

A.2.1 Summary Statistics Sample

Table A.1: Sample Characteristics (Main Experiment)

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	235	21.16	1.92	18	27
Female	234	0.54	0.50	0	1
Married	236	0.08	0.27	0	1
Occupation					
Student	236	0.83	0.38	0	1
Unemployed	236	0.10	0.30	0	1
Other	236	0.07	0.25	0	1
Estratificación socioeconómica					
1	234	0.05	0.21	0	1
2	234	0.38	0.49	0	1
3	234	0.47	0.50	0	1
4	234	0.09	0.28	0	1
5	234	0.02	0.13	0	1
6	234	0.00	0.07	0	1
Education					
High school	236	0.10	0.30	0	1
Some college semester	236	0.55	0.50	0	1
Technical degree (Técnico)	236	0.15	0.36	0	1
University degree	236	0.19	0.39	0	1

Notes: Table shows summary statistics for the participants' characteristics.

Table A.2: Sample Characteristics (Second Experiment)

Variable	Obs	Mean	Std. Dev.	Min	Max
Demographics					
Age	66	22.39	2.77	18	31
Female	66	0.47	0.50	0	1
Married	66	0.09	0.29	0	1
Occupation					
Student	66	0.91	0.29	0	1
Unemployed	66	0.05	0.21	0	1
Other	66	0.05	0.21	0	1
Estratificación socioeconómica					
1	66	0.03	0.17	0	1
2	66	0.42	0.50	0	1
3	66	0.42	0.50	0	1
4	66	0.09	0.29	0	1
5	66	0.03	0.17	0	1
6	66	0.00	0.00	0	0
Education					
High school	66	0.05	0.21	0	1
Some college semester	66	0.71	0.46	0	1
Technical degree (Técnico)	66	0.06	0.24	0	1
University	66	0.18	0.39	0	1

Notes: Table shows summary statistics for the participants' characteristics.

A.2.2 Treatment Balance

Table A.3: Balance Table (Main Experiment)

	Control	Arbitrary	Productivity	Needs	Double	Total	F-test
Age	21.711 (2.205)	20.446 (1.548)	21.288 (1.894)	21.321 (1.879)	21.269 (2.070)	21.162 (1.921)	0.008
Female	0.590 (0.498)	0.518 (0.504)	0.627 (0.488)	0.491 (0.505)	0.423 (0.504)	0.540 (0.499)	0.368
Married	0.128 (0.339)	0.054 (0.227)	0.085 (0.281)	0.071 (0.260)	0.077 (0.272)	0.081 (0.273)	0.818
Estrato	2.526 (0.687)	2.582 (0.896)	2.678 (0.753)	2.673 (0.840)	2.962 (0.774)	2.661 (0.805)	0.208
Degree	0.359 (0.486)	0.393 (0.493)	0.322 (0.471)	0.304 (0.464)	0.269 (0.452)	0.335 (0.473)	0.787

Notes: The table displays background characteristics (first column) for each treatment condition (first row) at the level of worker pairs and for the total sample. On average, workers are slightly younger in the Arbitrary treatment condition, significant according to an F-test, although the mean difference to the other conditions is 1.25 years at the maximum. "Degree" is a dummy for having a higher-education degree (Technical "Técnico" or University degree). The p-value we report in the last column is from an F-test of joint significance in a regression of the background characteristic on treatment indicators.

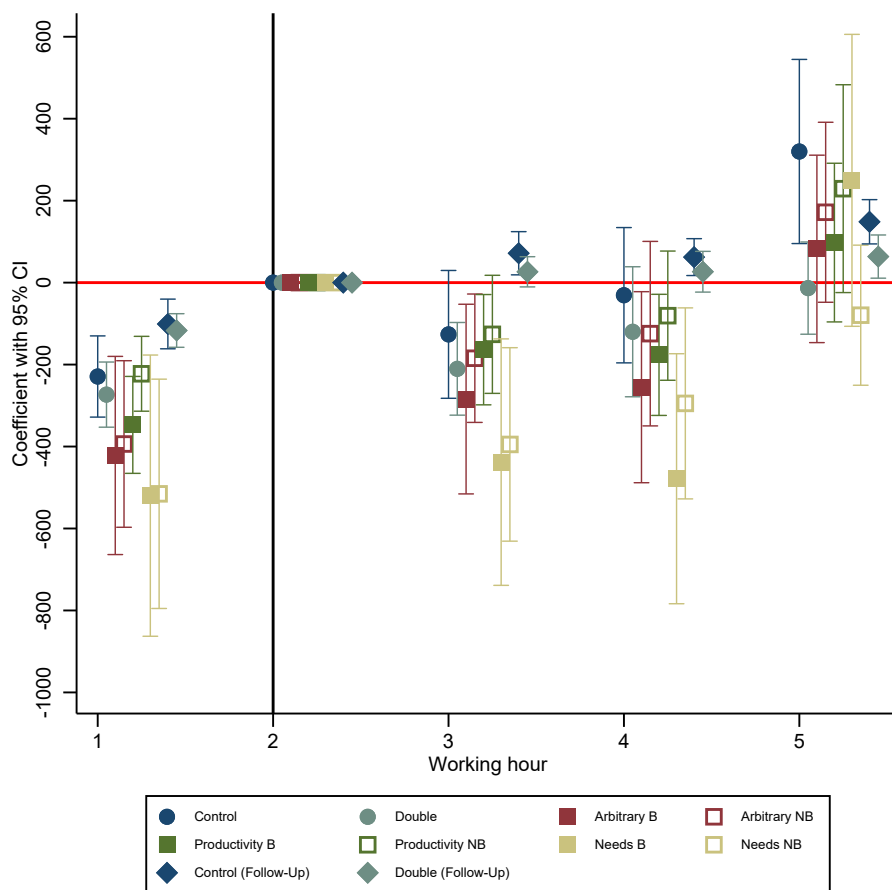
Table A.4: Balance Table (Second Experiment)

	Control	Double	Total	F-test
Age	22.364 (2.434)	22.424 (3.103)	22.394 (2.767)	0.930
Female	0.455 (0.506)	0.485 (0.508)	0.470 (0.503)	0.809
Married	0.030 (0.174)	0.152 (0.364)	0.091 (0.290)	0.089
Estrato	2.636 (0.929)	2.697 (0.684)	2.667 (0.810)	0.764
Degree	0.273 (0.452)	0.212 (0.415)	0.242 (0.432)	0.573

Notes: The table displays background characteristics (first column) for both treatment conditions (first row) at the level of worker pairs and for the total sample. "Degree" is a dummy for having a higher-education degree (Technical "Técnico" or University degree). The p-value we report in the last column is from an F-test of joint significance in a regression of the background characteristic on treatment indicators.

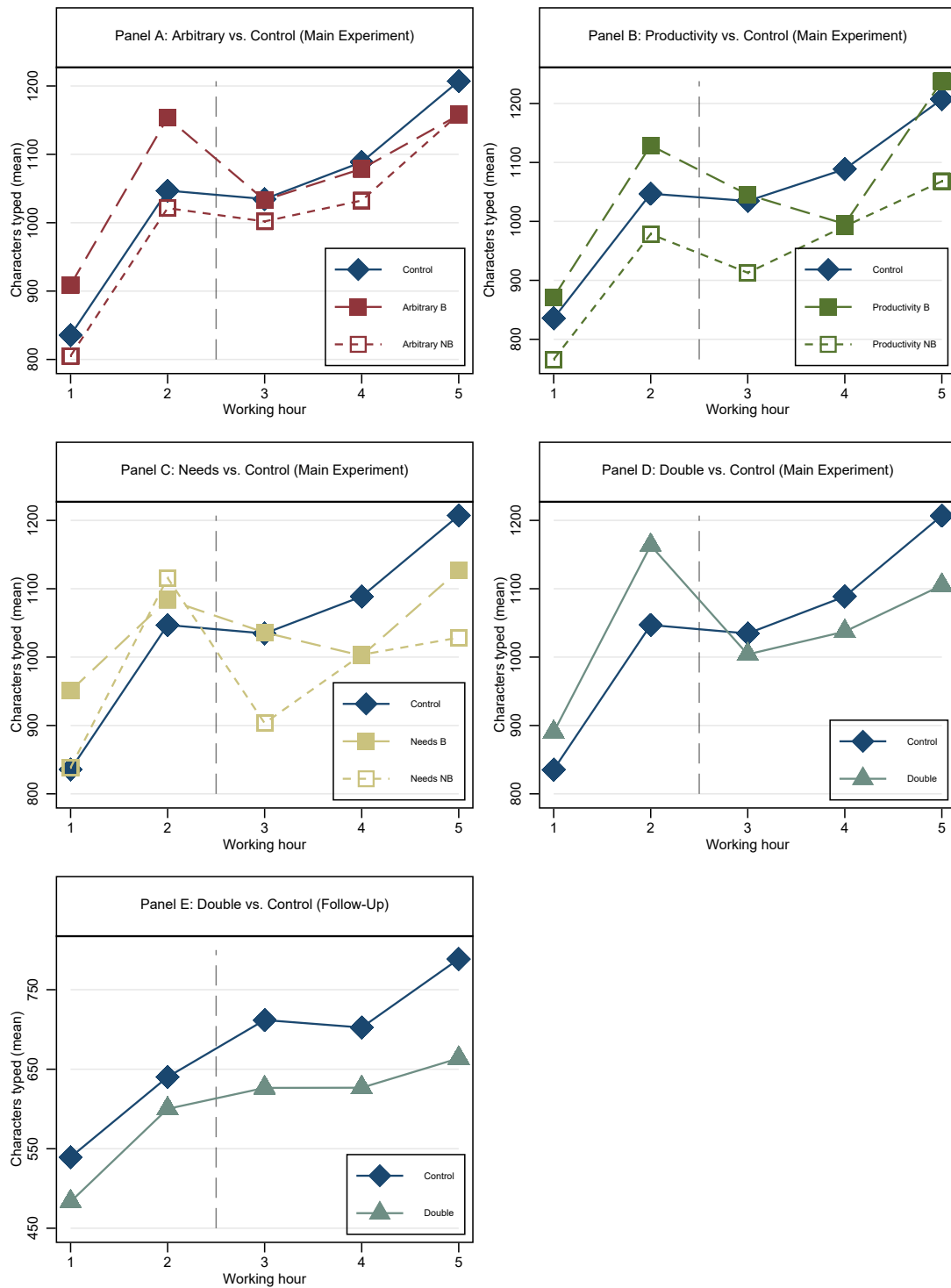
A.2.3 Supplementary Figures

Figure A.1: Event Study Plots (Both Experiments)



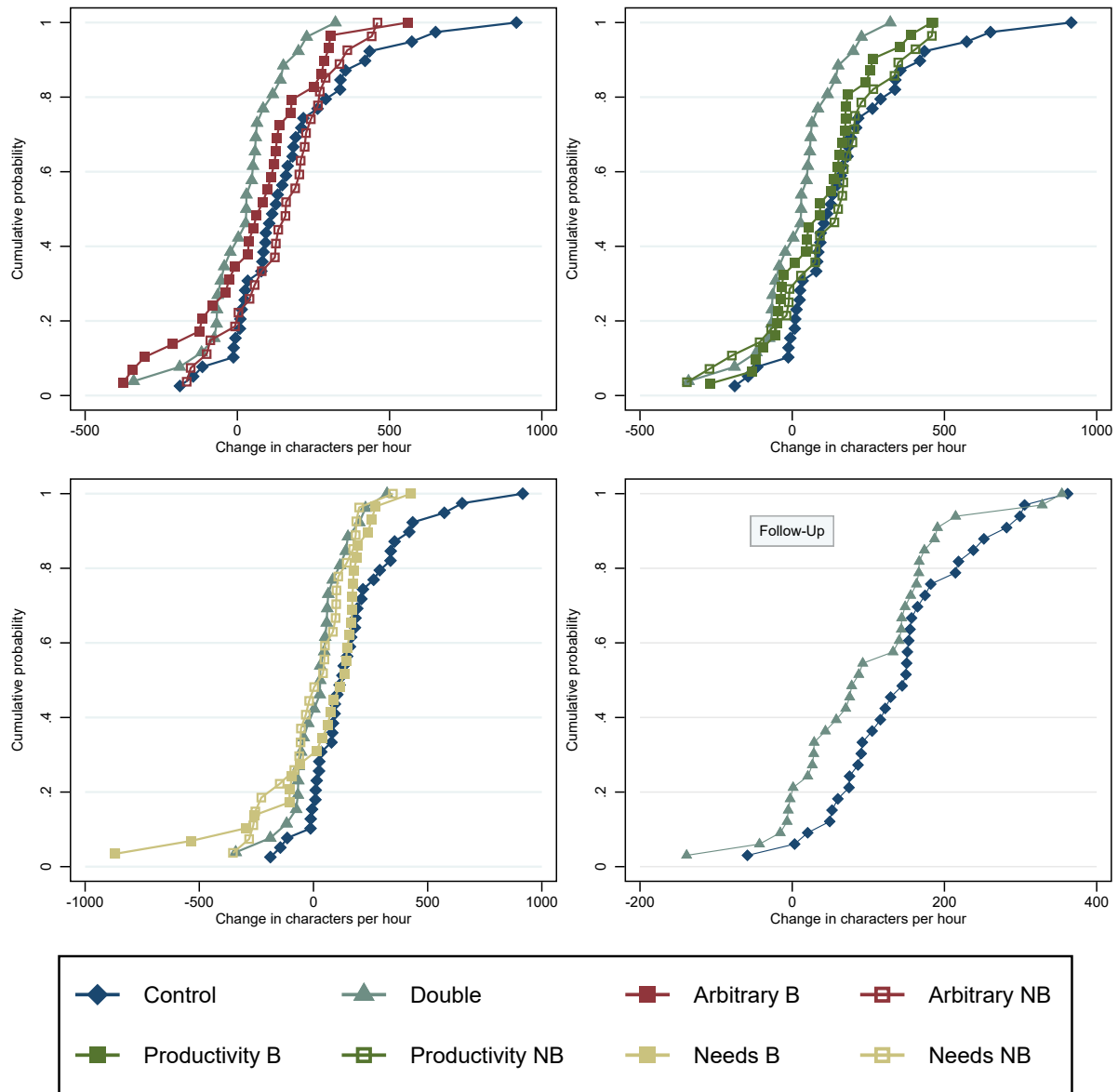
Notes: The figure plots event-study coefficients from a panel data regression with individual fixed effects. The dependent variable is the number of characters typed. Standard errors are clustered at the individual level. The regressions use the second working hour (the period immediately preceding treatment) as the omitted baseline category. The figure was generated using the eventdd command by D. Clarke and K. T. Schythe.

Figure A.2: Hourly Data: Typed Characters By Treatment Condition



Notes: The average number of typed characters is plotted for the five hours of the experiment for each treatment condition. The Control condition is plotted in every graph for reasons of comparison. If our assistants could not track hourly data for one hour due to technical reasons, they assigned the output count to the next hour. This practice leads to virtual productivity peaks. Due to this, the above figure shows "smoothed" data in which the average productivity of both hours was assigned to each of the hours in these cases. The figure looks qualitatively identical when plotting only observations for which the output was recorded for every single hour. Productivity dropped between sessions - likely due to a "siesta effect" after lunch - but this decline was much larger in all bonus treatments than in the Control, especially among bonus recipients. In the Needs treatment, productivity remained persistently lower after the break.

Figure A.3: ECDFs: Productivity Changes Between Morning and Afternoon Session (Both Experiments)



Notes: This figure displays the empirical distribution function of changes in productivity, measured as the difference in average output (characters per hour) between the morning and afternoon sessions. The plotted conditions correspond to those listed in the legend. The plot in the lower right corner presents results from the follow-up experiment.

A.2.4 Regressions with Afternoon Productivity as Dependent Variable

Table A.5 reports regression results using an ANCOVA-type specification where the dependent variable is average afternoon productivity (measured as characters typed per hour), and the independent variables are treatment dummies, with the control treatment omitted. The regression controls for productivity during the two hours of the morning session. Specifically, we estimate the following regression model, where β_0 is the intercept, *Double*, *Advantaged_T*, and *Disadvantaged_T* are treatment indicators, and *characters_{ik}* represents the number of typed characters by individual i in hour $k = \{1, 2\}$, included as controls for morning productivity. ϵ_i is the error term. Bootstrapped standard errors are clustered at the individual level to account for within-subject correlation.

$$prod_i = \beta_0 + \beta_1 \cdot Double + \sum_T \beta_{2_T} \cdot Advantaged_T + \sum_T \beta_{3_T} \cdot Disadvantaged_T + \sum_{k=1}^2 \alpha_{ik} \cdot characters_{ik} + \epsilon_i$$

The findings based on the ANCOVA-type regressions in Table A.5 confirm our main result obtained in the difference-in-differences specification: a significant decline in productivity following the introduction of a bonus payment, most strongly pronounced in the Double Bonus condition.

Table A.5: OLS Regressions: Afternoon Productivity as Dependent Variable

Dependent variable	Characters (1)	Characters (2)	Characters (3)	Characters (4)	Characters (5)	Characters (6)
Double	-97.466**** (25.82)	-128.831*** (45.67)	-48.320** (24.17)	-86.225**** (23.80)	-114.463*** (40.90)	-48.320** (24.17)
Arbitrary B	-59.505 (43.04)	-91.405* (49.45)		-57.787 (41.95)	-78.033 (47.72)	
Arbitrary NB	14.503 (38.07)	-23.622 (45.99)		19.844 (37.24)	-4.307 (43.18)	
Productivity B	-30.174 (34.62)	-63.654 (41.92)		-27.081 (33.39)	-48.363 (38.78)	
Productivity NB	-24.494 (43.63)	-64.628 (51.75)		-16.89 (42.28)	-42.41 (48.58)	
Needs B	-82.537* (44.99)	-115.263** (52.16)		-52.75 (38.00)	-74.486* (44.51)	
Needs NB	-125.104**** (37.95)	-159.962**** (44.43)		-122.386*** (37.71)	-144.415**** (43.39)	
Constant	280.393**** (38.64)	366.498**** (67.53)	132.307**** (32.96)	245.934**** (30.76)	300.304**** (53.93)	132.307**** (32.96)
Morning-session productivity	yes	yes	yes	yes	yes	yes
No. Individuals	302	236	66	300	234	66
Hypothesis tests (p-values)						
Double = Control	0.0002	0.0048	0.0456	0.0003	0.0051	0.0456
Double = Arbitrary B	0.3691	0.4503	.	0.4980	0.4569	.
Double = Productivity B	0.0483	0.1152	.	0.0771	0.0948	.
Double = Needs B	0.7398	0.8008	.	0.3782	0.3770	.
Arbitrary B = Productivity B	0.5245	0.5455	.	0.5041	0.5240	.
Arbitrary B = Needs B	0.6869	0.6752	.	0.9201	0.9450	.
Needs B = Productivity B	0.3058	0.3055	.	0.552	0.5476	.
Arbitrary NB = Productivity NB	0.4405	0.4190	.	0.4667	0.4473	.
Arbitrary NB = Needs NB	0.0018	0.0022	.	0.0017	0.0018	.
Needs NB = Productivity NB	0.0450	0.0595	.	0.0376	0.0447	.

Notes: This table shows OLS regression results. The dependent variable is the number of typed characters per hour in the afternoon-session. The first column presents the combined results from both experiments, while the second and third columns report the results of the first and second experiments separately. The fourth, fifth, and sixth column show regression results when excluding observations whose change in effort lies outside of Tukey's fences (below or above the first or third quartile by 1.5 times the interquartile range, respectively). Bootstrapped standard errors (10,000 replications) in parentheses clustered at individual level. **** p < 0.001, *** p < 0.01, ** p < 0.05, * p < 0.1..

A.2.5 Multiple Hypothesis Testing Adjustments

Because experiments often involve comparing outcomes across several treatment conditions with each other or relative to the relevant control condition, reporting p-values using methods controlling the familywise error rate should become the best practice in experimental work (List, Shaikh and Xu, 2019). The following tables show unadjusted p-values, Romano-Wolf p-values using the step-down procedure from the *rwolf2* STATA package by Clarke (2021) building on Romano and Wolf (2005), and the classical Bonferroni-Holm p-values. Standard errors are clustered at the individual level and based on 5,000 bootstrap replications using the *rwolf2* package. As a result, unadjusted p-values may differ slightly from those reported in the main text in some cases. Table A.6 presents the results from multiple hypothesis testing (MHT) for the main experiment based on the fixed-effects difference-in-differences estimation. The following Table A.7 shows corresponding MHT results for the ANCOVA-type regression analysis, where afternoon productivity is the dependent variable.

Table A.6: p-values with Multiple Testing Adjustments: Main Experiment

	Differences	Multiple testing adjustment		
		Unadjusted p-values	Romano-Wolf (<i>rwolf2</i>) p-values	Bonferroni-Holm p-values
Double vs. Control	-146.8	0.001	0.002	0.011
Disadvantaged vs. Control	-78.1	0.054	0.226	0.596
Advantaged vs. Double	105.2	0.229	0.718	1.000
Arbitrary B = Productivity B	-34.3	0.476	0.918	1.000
Arbitrary B = Needs B	21.0	0.733	0.956	1.000
Productivity B = Needs B	55.3	0.325	0.844	1.000
Arbitrary NB = Productivity NB	32.4	0.512	0.918	1.000
Arbitrary NB = Needs NB	149.8	0.001	0.003	0.014
Productivity NB = Needs NB	117.4	0.020	0.076	0.235
Arbitrary NB = Control	-17.7	0.705	0.956	1.000
Productivity NB = Control	-50.1	0.328	0.844	1.000
Needs NB = Control	-167.5	0.001	0.002	0.008
Arbitrary B = Double	36.5	0.433	0.910	1.000
Productivity B = Double	70.8	0.071	0.286	0.708
Needs B = Double	15.5	0.777	0.956	0.777

Notes: The table reports unadjusted and adjusted p-values from our hypotheses using data from the main experiment (236 workers). Differences refer to the pairwise differences across treatments regarding the difference in characters entered between the morning and the afternoon session. Romano-Wolf p-values with 5000 bootstrapping replications using standard errors clustered at the individual level. $k = 15$ is the number of unadjusted p-values under consideration. k_i is the number of p-values among the k p-values at least as large as the unadjusted p_i . The Bonferroni-Holm p-values are defined as $\min(1, k_i p_i)$. Romano-Wolf p-values calculated with the STATA package *rwolf2* (Clarke, 2021) using the step-down procedure.

Table A.7: p-values with Multiple Testing Adjustments: Main Experiment (ANCOVA-type)

	Multiple testing adjustment		
	Unadjusted p-values	Romano-Wolf (<i>rwolf2</i>) p-values	Bonferroni-Holm p-values
Double vs. Control	0.005	0.009	0.062
Disadvantaged vs. Control	0.036	0.076	0.427
Advantaged vs. Double	0.287	0.629	1.000
Arbitrary B = Productivity B	0.415	0.850	1.000
Arbitrary B = Needs B	0.675	0.855	1.000
Productivity B = Needs B	0.303	0.645	1.000
Arbitrary NB = Productivity NB	0.421	0.793	1.000
Arbitrary NB = Needs NB	0.003	0.005	0.036
Productivity NB = Needs NB	0.062	0.135	0.680
Arbitrary NB = Control	0.609	0.855	1.000
Productivity NB = Control	0.214	0.500	1.000
Needs NB = Control	0.0004	0.001	0.006
Arbitrary B = Double	0.448	0.793	1.000
Productivity B = Double	0.109	0.246	1.000
Needs B = Double	0.800	0.855	0.800

Notes: The table reports unadjusted and adjusted p-values from our hypotheses using data from the main experiment (236 workers) based on the ANCOVA-type analysis.. Romano-Wolf p-values with 5000 bootstrapping replications using standard errors clustered at the individual level. $k = 15$ is the number of unadjusted p-values under consideration. k_i is the number of p-values among the k p-values at least as large as the unadjusted p_i . The Bonferroni-Holm p-values are defined as $\min(1, k_i p_i)$. Romano-Wolf p-values calculated with the STATA package *rwolf2* (Clarke, 2021) using the step-down procedure.

A.2.6 Impact of Procedural Fairness in The Main Experiment

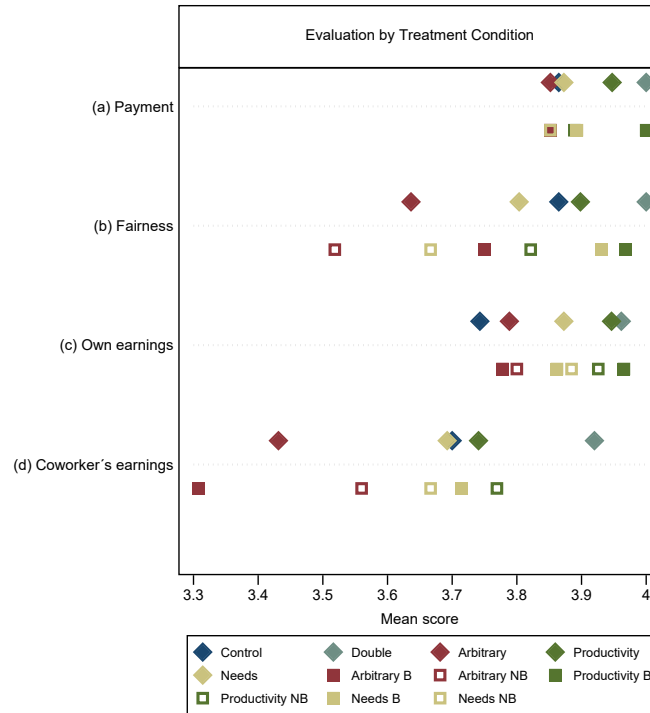
Most workers perceived the payment as adequate, with only minor differences between treatments (see Online Appendix Table A.9). Non-recipients in the Arbitrary and the Needs treatment report the lowest scores, in case of the latter significantly lower than in the Double Bonus condition ($p = 0.036$, Wald test).

Remarkably, all Double Bonus workers reported being treated fair, constituting the condition with the highest perceived fairness. Non-recipients in the Arbitrary treatment stand out for reporting the lowest average score for the overall fairness of how the workers were treated. Their evaluation score is marginally significantly lower than under the Control condition ($p = 0.074$, Wald test) and the Double Bonus condition ($p = 0.006$, Wald test). It is also at the margin of being significantly lower relative to non-recipients under the Productivity treatment ($p = 0.121$, Wald test; Table A.9, column 2). Considering bonus recipients and non-recipients together, reported fairness in the Arbitrary treatment is significantly lower than in the Productivity treatment ($p = 0.025$, Wald test) and the Double Bonus condition ($p = 0.001$, Wald-test; Table A.8, column 2).

Perception of deservedness of their own earnings in the afternoon session was highest for bonus recipients in the Productivity and the Double Bonus condition, reaching (marginal) statistical significance relative to the Control condition ($p = 0.074$ and $p = 0.081$, respectively Wald tests; Table A.9, column 4). The most sizable differences across treatments are observed when participants rated coworkers' earnings deservedness. Again, workers in the Arbitrary treatments expressed the lowest scores. Non-recipients in the Arbitrary treatment report a lower score than workers in the Double Bonus condition ($p = 0.050$, Wald test). Bonus recipients' score was even lower than non-recipients' score within the Arbitrary treatment and was lower than in the Control condition ($p = 0.110$, Wald test), the Double Bonus condition ($p = 0.003$, Wald test), and compared to bonus recipients in the Productivity treatment ($p = 0.071$, Wald test) and the Needs treatment ($p = 0.063$, Wald test). Bonus recipients and non-recipients in the Needs treatment rated the deservedness of coworker's earnings lower than workers under the Double Bonus condition ($p = 0.046$ and $p = 0.048$, respectively, Wald tests).

However, these mild differences in self-reported fairness perceptions in both Arbitrary treatment conditions failed to result in significant behavioral responses, i.e., productivity differences (see Table 2). More precisely, we detect no difference in effort reactions between Productivity and Arbitrary treatments neither for bonus recipients ($p = 0.479$, Wald test) nor for non-recipients ($p = 0.515$, Wald test). Hence, our data do not yield robust support for Hypothesis 4b. Bonus recipients' output increased more in the Productivity treatment than the Needs treatment, although the difference is insignificant at conventional levels (+5.7%, +55.3 characters per hour, $p = 0.335$, Wald test). Contrary to our expectations, non-recipients' output dropped in the Needs treatment significantly relative to both the Arbitrary (-15.4%, -149.8 characters per hour, $p = 0.001$, Wald test) and the Productivity (-12.1%, -117.4 characters per hour, $p = 0.020$, Wald test) treatment.

Figure A.4: Post-experimental Evaluation (Main Experiment)



Notes: This figure displays the average response scores from the post-experimental survey, grouped by treatment condition. The categories "Arbitrary," "Productivity," and "Needs" represent combined averages for both recipients and non-recipients within each respective condition. Answers could range between 1 "absolutely disagree" and 4 "absolutely agree" with respect to the following statements: (a) The payment I received was adequate. (b) The treatment of the two persons hired was fair. (c) The earnings I received in the second part of the day were deserved. (d) The earnings my coworker received in the second part of the day were deserved.

Table A.8: Regressions: Satisfaction Questions (Pooled Treatment Conditions, Main Experiment)

	(1) Payment Adequate	(2) Treatment Fair	(3) Own Earnings Deserved	(4) Other's Earnings Deserved
Double	0.135 (0.085)	0.135 (0.085)	0.219* (0.124)	0.220 (0.150)
Arbitrary	-0.013 (0.115)	-0.229* (0.136)	0.046 (0.141)	-0.269 (0.195)
Productivity	0.083 (0.090)	0.033 (0.096)	0.204* (0.121)	0.041 (0.154)
Needs	0.008 (0.098)	-0.061 (0.106)	0.130 (0.128)	-0.008 (0.152)
Constant	3.865**** (0.085)	3.865**** (0.085)	3.743**** (0.118)	3.700**** (0.140)
Obs.	228	233	224	212
R2	0.0145	0.0447	0.0337	0.0527
Hypothesis tests (p-values)				
Double = Control	0.113	0.114	0.079	0.144
Arbitrary = Control	0.910	0.095	0.748	0.171
Arbitrary = Double	0.062	0.001	0.048	0.001
Arbitrary = Productivity	0.256	0.025	0.061	0.041
Arbitrary = Needs	0.822	0.181	0.366	0.079
Productivity = Control	0.358	0.728	0.096	0.792
Productivity = Double	0.072	0.024	0.751	0.031
Productivity = Needs	0.195	0.228	0.208	0.575
Needs = Control	0.936	0.565	0.312	0.960
Needs = Double	0.011	0.003	0.158	0.004

Notes: The table shows OLS regression results from post-experimental questionnaire. Data from the first experiment. P-values for pairwise tests between treatment conditions reported. Standard errors clustered at session day level in parentheses. Column (1)-(4) have the answer score (1 absolutely disagree - 4 absolutely agree) to the following statements as the dependent variable. (1) The payment I received was adequate. (2) The treatment of the two persons hired was fair. (3) The earnings I received in the second part of the day were deserved. (4) The earnings received by my coworker in the second part of the day were deserved.

Table A.9: Regressions: Satisfaction Questions (Main Experiment)

	(1) Payment Adequate	(2) Treatment Fair	(3) Own Earnings Deserved	(4) Other's Earnings Deserved
Double	0.135 (0.085)	0.135 (0.085)	0.219* (0.124)	0.220 (0.151)
Arbitrary $\times$ NB	-0.013 (0.144)	-0.346* (0.192)	0.057 (0.155)	-0.140 (0.224)
Arbitrary $\times$ B	-0.013 (0.144)	-0.115 (0.158)	0.035 (0.171)	-0.392 (0.244)
Productivity $\times$ NB	0.024 (0.105)	-0.043 (0.124)	0.183 (0.129)	0.069 (0.165)
Productivity $\times$ B	0.135 (0.085)	0.103 (0.091)	0.223* (0.124)	0.0143 (0.174)
Needs $\times$ NB	-0.013 (0.110)	-0.198 (0.16)	0.142 (0.146)	-0.033 (0.182)
Needs $\times$ B	0.028 (0.116)	0.066 (0.098)	0.119 (0.135)	0.014 (0.166)
Constant	3.865**** (0.085)	3.865**** (0.085)	3.743**** (0.119)	3.700**** (0.141)
Obs.	228	233	224	212
R2	0.019	0.074	0.035	0.062
Hypothesis tests (p-values)				
Double = Control	0.115	0.116	0.081	0.147
Arbitrary NB = Control	0.928	0.074	0.713	0.534
Productivity NB = Control	0.820	0.727	0.159	0.675
Needs NB = Control	0.906	0.207	0.333	0.855
Arbitrary B = Control	0.928	0.468	0.839	0.110
Productivity B = Control	0.115	0.262	0.074	0.935
Needs B = Control	0.810	0.500	0.380	0.932
Arbitrary NB = Double	0.203	0.006	0.133	0.050
Productivity NB = Double	0.074	0.050	0.577	0.132
Needs NB = Double	0.036	0.012	0.408	0.048
Double = Arbitrary B	0.203	0.062	0.157	0.003
Double = Productivity B	1.000	0.321	0.938	0.074
Double = Needs B	0.176	0.153	0.189	0.046
Arbitrary B = Productivity B	0.203	0.113	0.146	0.071
Arbitrary B = Needs B	0.770	0.201	0.547	0.063
Needs B = Productivity B	0.176	0.527	0.164	1.000
Arbitrary NB = Productivity NB	0.778	0.121	0.265	0.282
Arbitrary NB = Needs NB	1.000	0.494	0.520	0.610
Needs NB = Productivity NB	0.691	0.332	0.677	0.474

Notes: The table shows OLS regression results from the post-experimental questionnaire. Data from the first experiment. P-values for pairwise tests between treatment conditions reported. Standard errors clustered at session day level in parentheses. Columns (1)-(4) have the answer score (1 absolutely disagree - 4 absolutely agree) to the following statements as the dependent variable. (1) The payment I received was adequate. (2) The treatment of the two persons hired was fair. (3) The earnings I received in the second part of the day were deserved. (4) The earnings received by my coworker in the second part of the day were deserved.

A.2.7 Analyses Based on The Questionnaire (Second Experiment)

Table A.10: Regressions: Questionnaire after Break (Second Experiment)

	(1) Satisfaction Work	(2) Employer Satisfaction	(3) Adequate Performance	(4) Firing Likely
Double	-0.029 (0.37)	0.036 (0.31)	0.112 (0.17)	-0.601* (0.32)
Married	0.368 (0.33)	0.517* (0.26)	-0.109 (0.27)	-0.038 (0.84)
Degree	-0.630 (0.65)	-0.140 (0.37)	-0.083 (0.20)	-0.660 (0.44)
Female	-0.475 (0.36)	-0.073 (0.31)	0.473*** (0.16)	0.295 (0.32)
Age	0.052 (0.08)	-0.016 (0.05)	0.018 (0.03)	-0.026 (0.07)
Estrato	-0.010 (0.18)	-0.055 (0.13)	-0.054 (0.13)	-0.093 (0.23)
Working day	-0.008 (0.02)	0.011 (0.01)	0.000 (0.01)	0.021 (0.02)
Confederate	0.180 (0.37)	0.178 (0.32)	-0.088 (0.22)	-0.270 (0.49)
Constant	5.753*** (1.82)	6.640*** (1.13)	5.837*** (0.78)	3.309* (1.68)
Obs.	66	66	66	66
R2	0.089	0.041	0.166	0.140
Hypothesis tests (p-values)				
Double = Control	0.938	0.909	0.507	0.067

Notes: The table shows OLS regression results from the questionnaire after the break and treatment administration in the second experiment. P-values for pairwise tests between treatment conditions reported. Standard errors clustered at session day level in parentheses. Columns (1)-(4) have answers to the following questions as dependent variable: (1) How satisfied are you with this job opportunity? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (2) How satisfied do you think your employer is with your work in these activities? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (3) How adequate do you think your performance in these activities has been? (Answer on 7-point Likert scale from Not at all adequate to Completely adequate), (4) How likely do you think it is that the employer hiring these activities will fire you for low productivity before the end of the day? (Answer on 7-point Likert scale from Totally unlikely to Totally likely). "Degree" is a dummy for having a higher-education degree (Technical "Técnico" or University degree). "Working day" is a linear time trend. "Confederate" is a dummy for the four cases where a recruited worker was paired with a confederate.

Table A.11: Regressions: Questionnaire After Break Without Control Variables (Second Experiment)

	(1) Satisfaction Work	(2) Employer Satisfaction	(3) Adequate Performance	(4) Firing Likely
Double	0.030 (0.35)	0.091 (0.26)	0.121 (0.16)	-0.545 (0.33)
Constant	6.394**** (0.26)	6.303**** (0.20)	6.273**** (0.11)	2.788**** (0.25)
Obs.	66	66	66	66
R2	0.000	0.002	0.010	0.039
Hypothesis tests				
Double = Control	0.931	0.731	0.459	0.104

Notes: The table shows OLS regression results from the questionnaire after the break and treatment administration in the second experiment. P-values for pairwise tests between treatment conditions reported. Standard errors clustered at session day level in parentheses. Columns (1)-(4) have answers to the following questions as dependent variable: (1) How satisfied are you with this job opportunity? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (2) How satisfied do you think your employer is with your work in these activities? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (3) How adequate do you think your performance in these activities has been? (Answer on 7-point Likert scale from Not at all adequate to Completely adequate), (4) How likely do you think it is that the employer hiring these activities will fire you for low productivity before the end of the day? (Answer on 7-point Likert scale from Totally unlikely to Totally likely).

Table A.12: Regressions: Compare After-Break and Final Questionnaire (Second Experiment)

	(1) Satisfaction Work	(2) Employer Satisfaction	(3) Adequate Performance
Double	0.030 (0.36)	0.091 (0.27)	0.121 (0.15)
Afternoon	0.242 (0.32)	0.121 (0.17)	-0.03 (0.10)
Double $\times$ Afternoon	0.121 (0.42)	-0.091 (0.28)	0.152 (0.13)
Constant	6.394**** (0.26)	6.303**** (0.19)	6.273**** (0.10)
Obs.	132	132	132
R2	0.019	0.003	0.033
Hypothesis tests (p-values)			
Double = Control (morning)	0.933	0.735	0.430
Double = Control (afternoon)	0.468	1.000	0.054
Afternoon	0.453	0.487	0.768
Double $\times$ Afternoon	0.772	0.747	0.258

Notes: The table shows OLS regression results for the questions asked both after the break and in the final questionnaire in the second experiment. P-values for pairwise tests between treatment conditions reported. Afternoon Standard errors clustered at individual level in parentheses. Columns (1)-(4) have answers to the following questions as dependent variable: (1) How satisfied are you with this job opportunity? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (2) How satisfied do you think your employer is with your work in these activities? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (3) How adequate do you think your performance in these activities has been? (Answer on 7-point Likert scale from Not at all adequate to Completely adequate). The interaction between "Double" and "Afternoon" as the difference-in-differences reveals no significant differences in the development of the outcome variables over the experiment's duration.

Table A.13: Likelihood of Getting Fired by Treatment Condition (Second Experiment)

	Control	Double	Total
Totally unlikely	4	9	13
Unlikely	16	16	32
Slightly unlikely	4	3	7
Neither unlikely nor likely	3	1	4
Slightly likely	5	4	9
Likely	0	0	0
Totally likely	1	0	1
Total	33	33	66
Wilcoxon rank-sum test	p = 0.088		

Notes: The table tabulates answers to the question "How likely do you think it is that the employer hiring these activities will fire you for low productivity before the end of the day?" asked after the break and treatment administration following the first session. p-value from a Wilcoxon rank-sum test against the null hypothesis of equal distributions reported.

Table A.14: Regressions: Final Questionnaire (Second Experiment)

	(1) Satisfaction Work	(2) Employer Satisfaction	(3) Adequate Performance	(4) Employer's Expectations	(5) Satisfaction Performance	(6) Importance Employer	(7) Compare Productivity	(8) Compare Effort	(9) Payment Adequate
Double	0.152 (0.21)	0.000 (0.11)	0.273* (0.15)	0.061 (0.24)	0.121 (0.14)	0.212 (0.24)	-0.03 (0.14)	0.061 (0.15)	-0.152 (0.20)
Constant	6.636**** (0.19)	6.424**** (0.07)	6.242**** (0.11)	5.182**** (0.17)	6.303**** (0.11)	5.848**** (0.17)	1.424**** (0.09)	1.545**** (0.10)	6.303**** (0.12)
Obs.	66	66	66	66	66	66	66	66	66
R2	0.008	0.000	0.057	0.001	0.014	0.014	0.001	0.002	0.007
Hypothesis tests									
Double = Control	0.476	1.000	0.070	0.803	0.393	0.373	0.826	0.683	0.453

Notes: The table shows OLS regression results for questions asked in the final questionnaire in the second experiment. P-values for pairwise tests between treatment conditions reported. Afternoon Standard errors clustered at individual level in parentheses. Columns (1)-(4) have answers to the following questions as dependent variable: (1) How satisfied are you with this job opportunity? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (2) How satisfied do you think your employer is with your work in these activities? (Answer on 7-point Likert scale from completely dissatisfied to completely satisfied), (3) How adequate do you think your performance in these activities has been? (Answer on 7-point Likert scale from Not at all adequate to Completely adequate), (4) How much do you believe your performance has met the expectations of the employer hiring for these activities? (Answer on 7-point Likert scale from Well below expectations to Well above expectations), (5) How satisfied are you with your performance in these activities? (Answer on 7-point Likert scale from very dissatisfied to very satisfied), (6) How important do you believe the work provided is, to the employer hiring for these activities? (Answer on a 7-point Likert scale from Not important to Extremely important), (7) How do you assess your hourly productivity (i.e., the total number of correctly transcribed entries divided by the 3 hours of work) in the second session, compared to the hourly productivity of the first session (i.e. e., the total number of correctly transcribed entries in the first session divided by the 2 hours of work)? (Answer on 3-point Likert scale from Higher to Lower in the second session than in the first session), (8) How do you evaluate your personal effort (on average per hour) in the second session, compared to the first session? (Answer on 3-point Likert scale from Higher to Lower in the second session than in the first session), (9) How adequate do you believe the payment you will receive for these activities is? (Answer on 7-point Likert scale from Not at all adequate to Totally adequate).

A.2.8 Additional Regressions Analyzing Productivity Changes

Table A.15 presents supplementary analyses from the first wave, using a simplified specification with two observations per individual - one in the morning and one in the afternoon. These include results using total entries, correct entries, and the share of correct entries (“Quality”) as outcome variables, as well as results for the main outcome variables obtained after excluding “severe” outliers identified using Tukey’s fences. Table A.16 presents supplementary results from the follow-up experiment, using entries as the dependent variable (columns 2 and 4) and excluding severe outliers (columns 3 and 4) following the methodology based on Tukey’s fences.

Table A.17 reports regressions using the same specification as in the main text, but with data pooled across both experimental waves. The findings confirm the significant post-bonus productivity decline when comparing the Double Bonus condition to the Control condition.

Table A.15: Additional Regressions (Main Experiment)

Dependent variable	Whole Sample			Outliers removed		
	(1) Entries	(2) Correct	(3) Quality	(4) Characters	(5) Entries	(6) Correct
Double $\times$ Afternoon	-5.885** (2.42)	-5.645** (2.37)	0.056 (0.55)	-127.101*** (39.48)	-4.676** (2.09)	-4.434** (2.04)
Arbitrary NB $\times$ Afternoon	-0.324 (2.57)	-0.151 (2.47)	0.020 (0.52)	1.94 (43.72)	0.885 (2.29)	1.06 (2.18)
Arbitrary B $\times$ Afternoon	-5.807** (2.91)	-5.506* (2.86)	0.182 (0.51)	-90.637* (49.06)	-4.598* (2.71)	-4.295 (2.66)
Productivity NB $\times$ Afternoon	-0.797 (3.09)	-0.936 (3.06)	0.023 (0.88)	-30.451 (48.44)	0.411 (2.85)	0.275 (2.82)
Productivity B $\times$ Afternoon	-3.658 (2.46)	-3.993 (2.45)	-0.805 (0.53)	-56.338 (41.57)	-2.449 (2.17)	-2.781 (2.15)
Needs NB $\times$ Afternoon	-9.151*** (2.87)	-8.713*** (2.78)	0.156 (0.53)	-147.819**** (44.75)	-7.942*** (2.67)	-7.501*** (2.56)
Needs B $\times$ Afternoon	-4.031 (2.83)	-3.805 (2.75)	0.022 (0.57)	-79.254* (48.16)	-2.868 (2.65)	-2.665 (2.55)
Afternoon	9.410**** (1.93)	8.799**** (1.87)	-0.699** (0.32)	149.171**** (29.75)	8.202**** (1.55)	7.588**** (1.47)
Constant	55.178**** (1.03)	54.138**** (1.03)	98.013**** (0.19)	968.748**** (17.56)	55.278**** (1.03)	54.235**** (1.02)
No. Individuals	236	236	236	234	234	234
Hypothesis tests (p-values)						
Double = Arbitrary B	0.976	0.957	0.832	0.441	0.977	0.958
Double = Productivity B	0.287	0.437	0.161	0.073	0.280	0.432
Double = Needs B	0.459	0.455	0.959	0.298	0.478	0.480
Arbitrary B = Productivity B	0.420	0.572	0.094	0.484	0.428	0.579
Arbitrary B = Needs B	0.553	0.564	0.799	0.835	0.577	0.593
Needs B = Productivity B	0.885	0.942	0.195	0.631	0.873	0.964
Arbitrary NB = Productivity NB	0.872	0.786	0.998	0.516	0.871	0.786
Arbitrary NB = Needs NB	0.001	0.001	0.820	0.001	0.001	0.001
Needs NB = Productivity NB	0.009	0.014	0.886	0.021	0.010	0.015

Notes: The table shows fixed effects regression results. In the first two columns, the dependent variables are the number of entries per hour and the number of correct entries per hour. In the third column, the dependent variable is the share of correct entries as a proportion of total entries. The remaining columns present results for characters, entries, and correct entries as the dependent variable, excluding observations with effort changes beyond Tukey's outer fences (i.e., below or above three times the interquartile range from the first or third quartile), thereby removing severe outliers, as also illustrated in the ECDFs above. Regressions are based on two observations (morning and afternoon average per hour) per individual. Bootstrapped standard errors (10,000 replications) in parentheses clustered at individual level. **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Hypothesis tests are Wald tests of the null hypotheses that the changes between morning and afternoon sessions in the dependent variable are equal in the contrasted conditions.

Table A.16: Detailed Regressions (Second Experiment)

Dependent variable	Characters (1)	Entries (2)	Characters (3)	Entries (4)
Double $\times$ Afternoon	-47.323** (23.68)	-2.717* (1.44)	-41.217** (21.00)	-2.482* (1.38)
Hour 2	108.848**** (17.62)	6.970**** (1.01)	102.778**** (17.78)	6.683**** (1.03)
Hour 3	181.525**** (17.97)	12.101**** (1.13)	172.678**** (17.38)	11.713**** (1.14)
Hour 4	176.919**** (20.31)	11.768**** (1.25)	167.599**** (18.81)	11.380**** (1.24)
Hour 5	238.510**** (23.21)	15.086**** (1.41)	227.202**** (21.30)	14.650**** (1.37)
Constant	511.364**** (22.52)	32.530**** (1.48)	513.746**** (23.36)	32.730**** (1.53)
No. Individuals	66	66	63	63
Hypothesis tests (p-values)				
Double = Control	0.046	0.059	0.050	0.071

Notes: This table shows fixed effects regression results for the second experiment. The dependent variable in the first two columns is the number of characters entered per hour. The dependent variables in the third and fourth column are the same variables when excluding outlier observations whose change in effort lies outside of Tukey's fences (below or above the first or third quartile by 1.5 times the interquartile range, respectively). Five observations (morning and afternoon) per individual. Bootstrapped standard errors (10,000 replications) in parentheses clustered at individual level. **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Hypothesis tests are Wald tests of the null hypotheses that the changes between morning and afternoon sessions in the dependent variable are equal in the contrasted conditions.

Table A.17: Regressions: Original Treatment Conditions (Data Pooled Across Both Experiments)

Dependent variable	Characters (1)	Entries (2)	Characters (3)	Entries (4)
Double $\times$ Afternoon	-93.602**** (25.58)	-4.104*** (1.43)	-82.915**** (23.66)	-3.458*** (1.30)
Arbitrary NB $\times$ Afternoon	-6.604 (38.01)	-0.365 (2.05)	4.083 (36.90)	0.281 (1.97)
Arbitrary B $\times$ Afternoon	-99.181** (43.77)	-5.848** (2.47)	-88.494** (41.87)	-5.202** (2.36)
Productivity NB $\times$ Afternoon	-38.995 (43.23)	-0.838 (2.65)	-28.308 (42.12)	-0.192 (2.59)
Productivity B $\times$ Afternoon	-64.882* (35.87)	-3.699* (1.91)	-54.195 (33.97)	-3.052* (1.78)
Needs NB $\times$ Afternoon	-156.363**** (38.71)	-9.192**** (2.43)	-145.676**** (37.16)	-8.546**** (2.35)
Needs B $\times$ Afternoon	-120.135** (52.66)	-4.072* (2.341)	-77.112* (40.70)	-3.472 (2.27)
Hour 2	304.868**** (27.49)	16.629**** (1.51)	308.887**** (26.94)	16.847**** (1.48)
Hour 3	195.702**** (25.45)	11.119**** (1.40)	185.980**** (23.62)	10.524**** (1.28)
Hour 4	230.960**** (27.18)	13.218**** (1.49)	221.887**** (25.71)	12.654**** (1.38)
Hour 5	503.785**** (42.90)	28.960**** (2.43)	496.547**** (42.87)	28.507**** (2.44)
Constant	730.427**** (18.63)	42.675**** (1.06)	725.653**** (18.37)	42.617**** (1.06)
No. Individuals	302	302	300	300
Hypothesis tests (p-values)				
Double = Control	0.0003	0.004	0.001	0.008
Double = Arbitrary B	0.893	0.457	0.893	0.456
Double = Productivity B	0.394	0.821	0.389	0.817
Double = Needs B	0.607	0.989	0.885	0.995
Arbitrary B = Productivity B	0.482	0.423	0.476	0.418
Arbitrary B = Needs B	0.736	0.552	0.831	0.566
Needs B = Productivity B	0.332	0.884	0.625	0.870
Arbitrary NB = Productivity NB	0.517	0.873	0.520	0.872
Arbitrary NB = Needs NB	0.001	0.001	0.001	0.001
Needs NB = Productivity NB	0.019	0.009	0.021	0.010

Notes: This table presents fixed effects regression results combining data from the main and second experiments. The dependent variable in the first and second columns is the number of typed characters and entries per hour, respectively. In the third and fourth columns, the dependent variables are the same but exclude observations with effort changes outside Tukey's fences (i.e., below or above 1.5 times the interquartile range from the first or third quartile). Five observations (morning and afternoon) per individual. Bootstrapped standard errors (10,000 replications) are shown in parentheses and are clustered at the individual level. Significance levels are indicated by **** $p < 0.001$, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Wald tests (below "hypothesis tests") examine the null hypothesis that changes in the dependent variable from morning to afternoon are equal across the contrasted conditions.

A.2.9 Treatment Effects Using Simple Statistical Tests

Table A.18: Treatment Effects by Treatment Condition (Main Experiment)

	Control	Arbitrary B	Arbitrary NB	Productivity B	Productivity NB	Needs B	Needs NB
Arbitrary B	-110.31 (-11.35%)						
p-value	0.093						
Arbitrary NB	-17.74 (-1.83%)	92.58 (9.53%)					
p-value	0.720	0.073					
Productivity B	-76.01 (-7.82%)	34.30 (3.53%)	-58.28 (-5.99%)				
p-value	0.191	0.631	0.163				
Productivity NB	-50.13 (-5.16%)	60.19 (6.19%)	-32.39 (-3.33%)	25.89 (2.66%)			
p-value	0.712	0.207	0.544	0.387			
Needs B	-131.27 (-13.51%)	-20.95 (-2.16%)	-113.53 (-11.69%)	-55.25 (-5.69%)	-81.14 (-8.35%)		
p-value	0.236	0.692	0.103	0.836	0.307		
Needs NB	-167.49 (-17.24%)	-57.18 (-5.89%)	-149.76 (-15.41%)	-91.48 (-9.42%)	-117.37 (-12.08%)	-36.23 (-3.73%)	
p-value	0.003	0.268	0.003	0.079	0.022	0.151	
Double	-146.78 (-15.11%)	-36.46 (-3.75%)	-129.04 (-13.28%)	-70.76 (-7.28%)	96.65 (9.95%)	-15.51 (-1.60%)	-20.72 (-2.13%)
p-value	0.003	0.288	0.004	0.089	0.024	0.117	0.887

Notes: The table presents treatment effects calculated as pairwise differences in differences in average productivity (typed characters per hour) from morning to afternoon (first experiment). Average treatment effects (ATEs) are shown as differences in differences (afternoon minus morning) between treatments, in both absolute terms and as percentages of the morning session average (971.53 characters per hour). Comparisons reflect the difference in the treatment condition specified in each column relative to the condition in each row. P-values are derived from two-sided Wilcoxon rank-sum tests, testing the null hypothesis of equal distributions between pairwise compared treatment groups.

Table A.19: Treatment Effects by Treatment Condition (Second Experiment)

	Control
Double	-47.32 (-8.4%)
p-value	0.057

Notes: The table presents treatment effects calculated as pairwise differences in differences in average productivity (typed characters per hour) from morning to afternoon (second experiment). Average treatment effects (ATEs) are shown as differences in differences (afternoon minus morning) between treatments, in both absolute terms and as percentages of the morning session average (565.79 characters per hour). Comparisons reflect the difference in the treatment condition specified in each column relative to the condition in each row. P-values are derived from two-sided Wilcoxon rank-sum tests, testing the null hypothesis of equal distributions between pairwise compared treatment groups.

Table A.20: Bonus Productivity Effects: Pooled Conditions

Condition	Diff-in-Diff	p-value (t-test)	p-value (rank-sum)	Cohen's d	Obs.
Pooled Experiments					
Double vs. Control	-93.60	0.0006	0.0006	-0.62	131
Bonus recipients vs. Non-recipients	-58.22	0.0060	0.0096	-0.32	302
First Experiment					
Double vs. Control	-146.78	0.0032	0.0026	-0.78	65
Bonus recipients vs. Non-recipients	-61.66	0.0186	0.0398	-0.31	236
Second Experiment					
Double vs. Control	-47.32	0.0535	0.0573	-0.48	66

Notes: The table reports treatment effects as pairwise differences-in-differences in average productivity (typed characters per hour) from morning to afternoon across pooled treatment conditions. The table reports p-values from two-sided t-tests and rank-sum tests, along with Cohen's d as a measure of effect size. Bonus recipients refers to all workers receiving a bonus, including those in the Double Bonus condition, while non-recipients includes all workers not receiving a bonus, including those in the Control condition.

A.3 Instructions and Experimental Protocol - Main Experiment

We report the experiment protocol and the instructions administered in the main experiment from 2014/2015 in Sections A.3.1 to A.3.5. All the text read aloud to participants is reported in italics. In order to ensure that the same information was administered to participants across sessions, the lead research assistant would read most of such instructions from a written script. Section A.3.6 reports further notes to the assistants. Section A.3.7 reports the instruction sheet given to participants to carry out their task. Section A.3.8 reports the final questionnaire. The original script in Spanish is available upon request.

A.3.1 Welcome

The lead research assistant (RA1 henceforth) (or the other research assistant, RA2 henceforth) greets the pair of workers at the Business School faculty or at the Konrad Lorenz University entrance. We try to avoid them talking with each other. We take one of them into the faculty lounge while the other sits waiting outside. We hand out the informed consent form, asking them to read it and to sign it. Tell them to please wait quietly before the day work starts. When they are both present and have signed the consent form, the schedule of the working day can be explained to them.

(At the Faculty reception desk or at the Assistant office): *Good morning guys, did you bring the copies of your electricity bills? Please give them to me.*

(If a participant does not bring them, he/she is asked to please call someone at home who can send them to him/her).

Before we begin, I am going to ask you to please read these informed consent form, which is to assure you that all the information you provide us today will be treated confidentially. If you have any questions, please let me know, otherwise please enter your ID and today's date and sign.

Once the informed consent is signed, the participants are asked to enter the staff room and sit in front of the computer where the task is explained.

Good morning again, thank you very much for your interest in collaborating with us. My name is [State name of lead researcher], and here with me is [State name of the other research assistant present].

RA2 acknowledges the introduction.

I am going to read the following instructions from a written text, because I do not want to forget important details. The work for which we are hiring you is part of a research project conducted and funded by a consortium of universities. We want to send some questionnaires to residents of Bogotá. For that purpose, we need to prepare the labels with the data of the recipients. Additionally, we want it to be a casual sample of Bogotá residents and so we will use a version of the telephone directory to extract the addresses of the residents. Both of you are going to do the same task individually. You will stay in two different rooms. Each of you will be responsible for the room and the computer

we give you, so we will ask you not to leave the room. If, for any reason, you do need to leave the room, I will give you shortly a phone number you can call. Laura/Daniel or I will come and stay in the rooms while you are out.

The work is divided into a two-hour session, a 15 minute break and then a three-hour session. The hourly rate is \$15,000, which means that at the end of the five hours you will receive \$75,000. Your contract is for today only. That is, under the terms of this project, you will not be offered a new contract. Do you agree with these conditions? Now I will show you your task for this job.

RA2 shows the Excel sheet ADDRESSES.xlsx.

On your computer you will find the file ADDRESSES, this is the file you have to fill in. This file has two columns, the City column, which will always be Bogotá because the directories are from Bogotá, and the Address column, which is the one you must fill in. Note that in the Address column, only the address goes, neither name nor telephone, if there is information of apartment, office, neighborhood, etc, include it, otherwise, only the address that appears. Now, where are you going to get the information to fill in the ADDRESSES file? From the COORDINATES file, which as its name indicates are search coordinates.

RA2 opens the second Excel sheet COORDINATES.xlsx.

As you can see, the file has the columns Page, Column and Entry, so in this first case for example, you search in your directory for page 93, column 4 and entry number 4.

RA2 shows an exercise.

Do you have any questions, is everything clear? Well, then to verify that it is clear, please do this example yourself.

RA2 invites the participants to execute an example.

If that is clear, there is only one exceptional case and it is this:

RA2 shows an example of entries that have no address.

Suppose you check your file, and the coordinate is this one that has no address associated with it, in that case then, you jump to the next address, and you write down the next available address, you do not leave blanks in the database, but you just write down the next address.

Well, finally I am going to ask you to please save each time you make an entry, either on the floppy disk icon or Control + G so as not to lose information.

Now, I am going to accompany Participant 1 to the room. I will be right back for you (Participant 2).

A.3.2 Morning Session

(In the offices of each participant):

Well, here I leave you your two files, I leave you water and in this envelope I leave you the instruction sheet with all the information I gave you a moment ago, with more detail, and the

phone number where you can call us for anything you need. If you don't have credit on your phone, you can use this cell phone. I leave you this envelope that has the example of how the labels we are going to print from the database you are helping us to build. So, it's XX:XX (state time); we will come round in two hours to let you know that the first session is over.

As soon as she closes the door, RA1 starts the timer and calls RA2 to activate his/her stopwatch. RA2 is present in the assistant office, but does not speak to the participant.

Repeat last set of operations for the second participant.

A.3.3 End of Morning Session

Every hour, RA2 saves the file in a folder and takes note of the number of completed addresses (and the number of characters) and does a random check of its accuracy/error rate. After exactly two hours, RA1 and RA2 go into the two offices separately.

Hi, the first 2-hour session is over. I brought you a snack, do you want to go to the bathroom or something? While you're out I'm going to save your work. I'm going to ask you to please stay away from the computer during these 15 minutes and exercise your hands to avoid fatigue. I will come back in 15 minutes for you to start the second session.

RA1 and RA2 take care that participants do not communicate with each other and go to the bathroom one at a time. They save the Excel file on the USB stick. RA2 computes relative productivity in Productivity Treatment; checks home quality as reported in the utility bill in Need Treatment; performs an unbiased random draw among the two workers in the Arbitrary treatment. It is then determined which participant will receive the bonus and who will not in the Treatment conditions. After 15 minutes, RA1 goes to each office and convenes the two participants to the staff room. RA2 is present in the teachers' room. When the two participants arrive, RA2 leaves the assistant office to watch the classrooms.

A.3.4 Treatment

The treatment sequence across sessions is the following:

1. Productivity Treatment
2. Need Treatment
3. Productivity Treatment
4. Arbitrary Treatment
5. Double Bonus Condition
6. Control Condition
7. Arbitrary Treatment
8. Needs Treatment
9. Control Condition
10. Double Bonus Condition

(In the Assistant room):

1. Control Condition: *We are ready to proceed with work. The work is the same as in the first session, but the duration of the second session will be 3 hours, and the wage will be the same as in the first part.*
2. Productivity Treatment: *We are ready to proceed with work. The work is the same as the first session, but the duration of the second session will be 3 hours. The wage for this afternoon will be \$15,000 per hour, as in the morning. However, the research director wants to pay a bonus of \$25,000 Pesos to one of you. To decide who would get the bonus, we reviewed the number of entries and characters you completed in the first session and, taking those criteria into account, you (indicating the participant receiving the bonus) will receive the bonus, which means we will pay you \$75,000 for the five hours plus a \$25,000 bonus for a total of \$100,000. In the meantime, you (indicating the participant not receiving the bonus) will receive the \$75,000 that we agreed upon this morning.*
3. Needs Treatment: *We are ready to proceed with work. The work is the same as the first session, but the duration of the second session will be 3 hours. The wage for this afternoon will be \$15,000 per hour, as in the morning. However, the research director wants to pay a bonus of 25,000 Pesos to one of the pair of workers. To decide who would get the bonus we reviewed the socio-economic strata of your houses, the poverty level of the locality where you live, and your age. Taking those criteria into account, you (indicating the participant receiving the bonus) will receive the bonus, which means that we will pay you \$75,000 for the five hours plus a \$25,000 bonus for a total of \$100,000. In the meantime, you (indicating the participant not receiving the bonus) will receive the \$75,000 that we agreed upon this morning.*
4. Arbitrary Treatment: *We are ready to proceed with work. The work is the same as the first session, but the duration of the second session will be 3 hours. The wage for this afternoon will be \$15,000 per hour, as in the morning. However, the research director wants to pay a bonus to one of you. Unfortunately, he did not have time to review who completed the most entries in the first session, or your socioeconomic status, age, or other criteria. He chose you (indicating the participant receiving the bonus) to receive the bonus, which means we will pay you \$75,000 for the five hours plus a \$25,000 bonus for a total of \$100,000. Meanwhile, you (indicating the participant receiving the bonus) will receive the \$75,000 that we agreed upon this morning.*
5. Double Bonus Condition: *We are ready to proceed with the work. The work is the same as the first session, but the duration of the second session will be 3 hours. The wage for this afternoon will be \$15,000 per hour, as in the morning. However, the research director wants to pay you both a bonus of 25,000 Pesos, which means we will pay you \$75,000 for the five hours plus a \$25,000 bonus for a total of \$100,000.*

Thank you. I will now take you to your offices.

RA1 separately takes the workers to their offices.

A.3.5 Afternoon session

As of now, the second session begins, it is XX:XX at XX:XX (state the time) and in three hours we will finish the second work session. Thank you.

Every hour, RA2 saves the file in a folder and takes note of the number of completed addresses (and the number of characters) and does a random check of its accuracy/error rate. After exactly three hours, RA2 goes to the first office:

The 3 hours are up. We would be grateful if you could answer this evaluation questionnaire. The questionnaire will be associated with a code so that your identity does not appear and the confidentiality of the information you provide is guaranteed. We will collect your questionnaire in this box.

Show ballot box, save all data in the USB flash drive and hand in the questionnaire.

I will be back in a few minutes with the money. In the meantime, please fill out the questionnaire.

RA2 goes to the other office and repeats. After 3-4 minutes she comes back to the first office.

Thank you very much for your cooperation. Please insert the questionnaire in the ballot box. Here is your money. Please confirm that it is complete (wait for the money to be counted). Now please fill out the receipt and I will take you to the elevator.

RA2 brings Worker 1 to the elevator, then repeats the process with Worker 2.

A.3.6 Further notes for assistants

Important tips when interacting with participants

1. Do not mention anything related to other workers.
2. Ask immediately for their CV (if for some reason this was not sent with the email) and the electricity bill.
3. Mention that it is a one-day job; there will be no possibility of contract extension.
4. Do not give any reference as to what we are expecting in terms of productivity.
5. Do not anticipate anything related to the bonus in the second session.
6. Always give the same answers to similar questions.
7. If you don't know what to say, stress that you are the project assistant and are only executing directives from the project coordinator; don't make anything up!
8. If a participant does not show up, find a university student willing to work.

Answers to Frequently Asked Questions

- Why do I have to fill out the consent form, and why the questionnaire?

The consent serves as a statement that you are aware of all procedures and the use of the data. The questionnaire is solely for our evaluation of working conditions and your satisfaction.

- What happens if I cannot complete all the addresses on my sheet?

Don't worry, nothing happens.

- I am entitled to an hour's break.

Since this is not a full workday, we would expect 15 minutes of rest to be sufficient. We also do this so that a student has more time to attend classes. However, if you need a longer break, we can accommodate it.

A.3.7 Task Instruction Sheet

1. Open the sheet COORDINATES SURNAME.xlsx.
2. Read the entries in the given order (starting with line 1, then 2, then 3 etc.).
3. The first number is the page number in the telephone directory.
4. The second number is the column number in the page (starting from the left).
5. The third number is the address line starting from the top. If the number is greater than the number of addresses, continue counting (same column) from top to bottom until you reach the number that appears in COORDINATES SURNAME.xlsx.

Figure A.5: Task Example

Guarín Torres Leonardo David	Av Jiménez 03-95 Candelaria	560 2943
Guarín García Luis Hernando	Cr 9 12-45 Int 1 Centro ...	560 2874
Guarín Hernández Luz Marina	Cr 65A 68A-62 La Estrada	542 6213
Guarín Segunda F. de	CI 72B 71B-10	251 9864
Guatame Ávila Andrés	CI 66 12-04 Chapinero	542 5863
Guatame Arévalo Marco Fidel	Cr 13 198-52 L-18 Canaima	528 1415
Guatava Ana Cecilia	Cr 13 61-05 L-292 Chapinero	542 0311
Guatava Ana Cecilia	Cr 13 61-05 L-292 Chapinero	542 0311

1. If the corresponding entry does not have an address but only a telephone number, please record the subsequent address, i.e. that in the next entry.
2. Open the sheet ADDRESSES SURNAME.xlsx.
3. Fill in the box in the "Address" column.
4. Fill in the "City" box by typing "Bogota".
5. Save.
6. Your room is reserved, if anyone asks something, answer that it is reserved by Professor Castiblanco from the Business School.
7. Remember that you are responsible for the computer and the classroom.
8. If you have any questions, please call immediately 314 3 06 XX XX XX.

A.3.8 Final Questionnaire

Code:

Q1) Sex: M F

Q2) Age: _____ years

Q3) What is your marital status?

- Married/Living with partner
- Single
- Separated/Divorced/Widow

Q4) According to your utility bills, what is the tier of your current home or neighborhood?

1	2	3	4	5	6
---	---	---	---	---	---

Q5) What is your level of education?

- None
- Primary School
- High school
- Some university, but no graduation
- Technical
- University degree

Q6) What is your father's level of education?

- None
- Primary School
- High school
- Some university, but no graduation
- Technical
- University degree

Q7) What is your mother's level of education?

- None
- Primary School
- High school
- Some university, but no graduation
- Technical
- University degree

Q8) What is your occupation?

- Student
- Unemployed, retired, housewife, househusband
- Other (specify) _____

Q9) For each of the statements below, please indicate your level of agreement or disagreement.

a. The work organization was effective.

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

b. The payment I received was appropriate.

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

c. The treatment received by the two people hired was equitable.

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

d. The earnings I received in the second part of the workday were well deserved.

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

e. The earnings that the other person hired in the second part of the workday received were well deserved.

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

Q10) Below are some possible reasons why some individuals succeed and others do not. On a scale of one to four, where one represents "not important" and five "very important"; indicate, in determining a person's success, how important it is:

a. Money inherited from family.

Not important	1	2	3	4	Very important
---------------	---	---	---	---	----------------

b. Hard work and initiative.

Not important	1	2	3	4	Very important
---------------	---	---	---	---	----------------

c. Connections and familiarity with the right people.

Not important	1	2	3	4	Very important
---------------	---	---	---	---	----------------

Q11) Among the following two factors, which factor do you consider to be more important for a person to be in a state of poverty?

- Lack of effort and work commitment on the part of the person.
- Luck or events that are not in the person's control

Q12) How much do you agree that the government needs to reduce the gap between rich and poor, either by raising taxes for the rich or by providing income assistance to the poorest? Please indicate how strongly you agree by marking a number from one to four on the scale below, where one indicates "strongly disagree" and four indicates "strongly agree".

Strongly disagree	1	2	3	4	Strongly agree
-------------------	---	---	---	---	----------------

Q13) How did you hear about this project? You can check more than one option.

- Through a poster
- By email
- Through social media
- Through a friend
- Through another person who was hired
- Other (specify) _____

Q14) Please express, if you wish,

your opinions about this workday. Thank you for your cooperation.

A.4 Instructions and Experimental Protocol - Second Experiment

The design of our follow-up experiment mirrors that of the original study. Notably, between 2014 and 2023, Colombia grappled with a turbulent period of economic instability, reflected in the stagnant GDP per capita, due to variable oil prices and internal challenges. The period was punctuated by a significant peace agreement with a major guerrilla group in 2016, fostering temporary stability, although pockets of rebellion persisted. Persistent inequality, unemployment, and rural underdevelopment compounded by an influx of refugees and the COVID-19 pandemic, created further challenges. Nevertheless, Colombia’s resilience, particularly in sectors like IT and communications, offer hope for economic recovery.

A notable divergence from the first experiment was the representation of the employer: due to funding constraints, we presented the employer as one of our researchers, a shift we highlighted in the pre-registration due to its potential impact on workers’ perceptions of employer authority. Accounting for eight years of inflation in line with annual adjustments to the Colombian minimum wage since the initial experiment, the revised advertised hourly wage was set at 27,000 Pesos (around 6.7 USD). Bonuses amounted up to 45,000 Pesos (about 10.8 USD), making up a third of the overall advertised earnings.

The cumulative inflation from 2015 to 2023 was 57%, while the increase in the minimum wage outpaced it at 80%. We decided to adjust our wages based on this latter indicator. Specifically, 27,000 is exactly 180% of 15,000, and 45,000 is 180% of 25,000. As a validity check, the Big Mac Index (NASDAQ Data Link: Big Mac Index Colombia, 2023) shows that the Big Mac price ranged from 7,900 to 8,600 Pesos in 2014-2015. By summer 2023, it had risen to 14,950 Pesos. Using the average initial Big Mac price of 8,250 Pesos for adjustment, the original 15,000 Pesos wage corresponds to roughly 27,000 Pesos in 2023 terms.

In 2023, the experiment sessions spanned from May 15th to July 7th at the Universidad Nacional de Colombia’s Bogotá campus. Sixty-six participants were engaged, with an even split between the Control and Double treatment groups. When a designated worker failed to attend, a confederate filled in. These confederates were not included in our analysis sample, and the recruited workers were unaware of working with a confederate. As a novelty compared to the initial experiment, we introduced several questions to probe more deeply into workers’ perceptions. These covered topics from their perceptions of employer satisfaction, views on their effort’s adequacy to the perceived risk of early dismissal. A concise questionnaire which was framed as a self-assessment, was presented post-treatment, between the morning and afternoon sessions. A detailed questionnaire followed the day’s conclusion, preceding the payment distribution.

In the subsequent section, we detail the experimental protocol and the guidelines implemented during our second experiment. Text presented in *italics* represents information directly communicated to the participants. We provide the original Spanish version of the script upon request.

A.4.1 Email

Good afternoon. We are writing to you because you expressed interest in collaborating on research activities and projects. A consortium of universities is seeking research assistants to perform certain tasks. The project is led by the researcher Dr. Grimalda from the Kiel Institute for the World Economy. The activity in question involves data entry. The hourly pay is 27,000 Pesos. This involves a single day, for five hours. If you are interested in this opportunity, please fill out this form (link) to confirm your availability. Punctuality at work is fundamental to us, so please keep that in mind when confirming your availability. As Excel knowledge is required, please attach your resume in the form. We will not keep the resumes after completing the process, whether the selection is positive or if the profile is discarded.

Sincerely, UEC CID

A.4.2 Second Email

Good afternoon. Continuing with the process for Dr. Grimalda's project, and given your availability, we invite you for the day ##/##/##, at 7:45. We recommend punctuality. The activity will take place in building A4, on the campus of the National University (in the Camilo Torres Area), here we put a link to Maps for the place.

Sincerely, UEC-CID

A.4.3 The day of the hiring

Checklist:

Room 803

- 1 computer where 2 example files DIRECCIONES.xlsx and COORDENADAS.xlsx are stored
- 1 telephone directory with white pages.
- Printed sheet of instructions.
- 1 envelope with an address printed on a label and attached.

Rooms 701 and 703

- 2 computers with DIRECCIONES.xlsx and COORDENADAS.xlsx files; The DIRECCIONES sheet is already saved with office number and date. The COORDENADAS.xlsx sheet is protected, so the participant cannot modify it. Both sheets are saved in Drive. [create two Google accounts and leave only the Excel files of the day, then download them and upload them to the Drive of the shared project folder]
- 2 telephone directories (2018 and 2016).
- 2 printed instruction sheets.
- 2 printed instruction sheets.

With us

- Mobile stopwatch
- Payment receipts.

The assistant has the first personalized meeting for each pair. Important tips:

- 1) Do not mention anything related to other participants.
- 2) Do not give any reference about what we are expecting in terms of productivity.
- 3) Do not anticipate anything related to the bonus of the second session.
- 4) Always give the same answers to similar questions.
- 5) If you do not know what to say, underline that you are the project assistant and you are only executing directives from the project coordinator; do not make anything up!
- 6) If a worker does not show up, call the confederate. The pair of workers (can be mixed or same sex, we do not mention anything in the ad) are received at the entrance of the CID. If possible, try to separate them so they do not talk before the investigation. Tell them to please wait in silence while the work day begins.

A.4.4 Act 1: Room 803

Good morning, thank you very much for your interest in collaborating with us.

The job for which we are hiring you is paid for by an international consortium of universities. The project is led by the researcher Dr. Grimalda from the Kiel Institute for the World Economy. We need to generate a random sample of residents in Bogota and for that we need to generate a random sample of addresses. For this reason, we will use a version of the white pages to extract addresses. You will do the same task individually. You will be in different offices. Each of you will be responsible for the office and the computer assigned to you, so it is advisable not to leave the office. If you need to leave the office to go to the bathroom, please lock it and, when you want to return, look for me here and I will open it again. The job is divided into a two hour session, a 15 minute break, and then a three hour session. Your task will end today, that is, in terms of this project, no new contract will be offered to you. We are going to pay you \$27000 Pesos per hour. You will receive your payments at the end of today's workday.

Verify that they have Nequi/Daviplata for payment.

We will now explain the activity to you.

A1 shows the DIRECCIONES.xlsx Excel sheet.

In this Excel sheet, you will find the fields Address, City. You will work with similar sheets on the computers that are set up in the two offices where you will work. We have already created an Excel sheet with the office number of each of you.

The assistant shows the procedure. Then opens the second Excel sheet COORDENADAS.xlsx

Now, in this other Excel sheet you will find a series of numbers. The sheet has three columns: one says page, the second says column, the third says entry. This sheet is protected, so it cannot be modified.

The assistant shows the telephone directory.

Now the procedure is as follows. You have to look in consecutive order at the rows of the COORDENADAS.xlsx sheet. The first one says 245, 3, and 20. This means that you have to open the directory on page 245 and look for the address that is in column three (from left to right), entry

20 (from These numbers were drawn randomly in advance, so in this way, we can create a random sample of Bogotanos. Be careful, we are talking about entries and not rows because an entry can occupy more than one row. In some cases, an entry will not have any address, but a phone number. In these cases, please put the address of the next entry in the same column. You do not have to put the name or surname of the entry, just the address because the guides are old. This will allow generating envelopes with labels like this (show the envelope).

Once you have identified the address, copy the data into your work sheets in the following way.

The assistant shows the task on the computer.

Is the task clear? Please, do a test now to make sure everything is clear.

RA invites each one to do a test, asking them to write the address in the corresponding box. Check that the solution is correct. If it is not, explain again and do another test. After both have completed a test correctly, continue.

Very good. You will have next to the computer a sheet with the instructions. Please check it now to make sure you have understood. If you have questions, please ask them now. I will be watching. After 2 hours I will go to your office and review your work. Are all the rules clear to This is your office. Here on this computer, the "addresses" and "coordinates" sheets are already open. Here you have a bottle of water. You can start now, I'll leave you to it. I will come here exactly in 2 hours. Remember that no one can enter this office. Please focus on the work, do not waste time and work individually in your office, for which you are responsible. Have a good work!

As soon as the door is closed RA starts the stopwatch.

A.4.5 End Act

Every hour, RA notes the number of addresses completed and the number of characters, downloading the file from the drive so it is not noticeable.

A.4.6 Act 2: After two hours

Hello, the first session has ended. Now rest for 15 minutes. I will return in 15 minutes to give you instructions on the second part. I brought you a snack. If you want to go to the bathroom, do it at once and I'll close the office. Please do not leave the office unattended. Repeat for each one.

A.4.7 Act 3

Room 803

The sequence is (you have to complete 15 rounds):

Control

Double

1. Control. We are ready to continue with the work. The job is the same as the first session, but the duration of the second session will be 3 hours. The pay will be 27,000 per hour as in the

first session. At the end of the day, we will pay you 135,000 Pesos. I will now accompany you to your office.

2. Double Condition. We are ready to continue with the work. The job is the same as the first session, but the duration of the second session will be 3 hours. The pay will be 27,000 per hour as in the first session. However, we want to pay a premium of 45,000 Pesos to each of the pair of workers. At the end of the day, we will pay you 180,000 Pesos. I will now accompany you to your office. RA accompanies the workers to their offices. Before starting the second session, we appreciate you filling out this self assessment exercise of your performance so far. The confidentiality of the information you provide us is guaranteed.

1. How satisfied are you with this job opportunity?
 - (a) Completely dissatisfied
 - (b) Dissatisfied
 - (c) Partially dissatisfied
 - (d) Neither dissatisfied nor satisfied
 - (e) Partially satisfied
 - (f) Satisfied
 - (g) Completely satisfied
2. How satisfied do you think your employer is with your work in these activities?
 - (a) Completely dissatisfied
 - (b) Dissatisfied
 - (c) Partially dissatisfied
 - (d) Neither dissatisfied nor satisfied
 - (e) Partially satisfied
 - (f) Satisfied
 - (g) Completely satisfied
3. How adequate do you think your performance in these activities has been?
 - (a) Not at all adequate
 - (b) Very little adequate
 - (c) Little adequate
 - (d) Somewhat adequate
 - (e) Moderately adequate
 - (f) Quite adequate
 - (g) Completely adequate
4. How likely do you think it is that the employer hiring these activities will fire you for low productivity before the end of the day?
 - (a) Totally unlikely
 - (b) Unlikely
 - (c) Somewhat unlikely
 - (d) Neither unlikely nor likely

- (e) Somewhat unlikely
- (f) Likely
- (g) Totally likely

You can start now, I'll leave, I'll be here exactly in 3 hours. Remember that no one can enter this office. Have a good job!

As soon as the door is closed, start timing. Repeat process for the second worker.

Every hour, RA takes note of the number of addresses completed and the number of characters.

After exactly three hours, RA goes to the first office:

The 5 hours are up. We appreciate you filling out this self assessment exercise of your performance, at the end of the workday. The confidentiality of the information you provide us is guaranteed. I'll be back in a few minutes. In the meantime, please fill out the questionnaire.

Assistant goes to the other office and repeats. After 3-4 minutes returns to the first office.

Thank you very much for your cooperation. Can you provide me your Nequi number for payment? Now please fill out the receipt and I'll accompany you.

The assistant brings Worker 1 to the exit, then repeats the process for Worker 2.

Perform a random check of entries for each worker and record the result.

A.4.8 Questionnaire

- Q1) How satisfied are you with this job opportunity?
 - Totally dissatisfied
 - Dissatisfied
 - Partially dissatisfied
 - Neither dissatisfied nor satisfied
 - Partially satisfied
 - Satisfied
 - Totally satisfied
- Q2) How satisfied do you think the employer hiring for these activities is with your work?
 - Totally dissatisfied
 - Dissatisfied
 - Partially dissatisfied
 - Neither dissatisfied nor satisfied
 - Partially satisfied
 - Satisfied
 - Totally satisfied
- Q3) How adequate do you believe your performance in these activities has been?
 - Not at all adequate
 - Very little adequate
 - Somewhat adequate
 - Fairly adequate

- Moderately adequate
 - Quite adequate
 - Totally adequate
- Q4) How much do you believe your performance has met the expectations of the employer hiring for these activities?
 - Well below expectations
 - Below expectations
 - Slightly below expectations
 - Meets expectations
 - Slightly above expectations
 - Above expectations
 - Well above expectations
- Q5) How satisfied are you with your performance in these activities?
 - Very dissatisfied
 - Dissatisfied
 - Somewhat dissatisfied
 - Neutral
 - Somewhat satisfied
 - Satisfied
 - Very satisfied
- Q5) How important do you believe the work provided is, to the employer hiring for these activities?
 - Not important
 - Little important
 - Somewhat important
 - Moderately important
 - Important
 - Very important
 - Extremely important
- Q6) How do you assess your hourly productivity (i.e., the total number of correctly transcribed entries divided by the 3 hours of work) in the second session, compared to the hourly productivity of the first session (i.e. e., the total number of correctly transcribed entries in the first session divided by the 2 hours of work)?
 - Higher in the second session than in the first session
 - Approximately equal in the second session than in the first session
 - Lower in the second session than in the first session
- Q7) How do you evaluate your personal effort (on average per hour) in the second session, compared to the first session?
 - Higher in the second session than in the first session

- Approximately equal in the second session than in the first session
 - Lower in the second session than in the first session
- Q8) Please indicate your level of agreement or disagreement with this statement: Even if they receive a bonus, workers do not have to make more than the standard effort.
 - Totally disagree
 - Disagree
 - Partially disagree
 - Neither disagree nor agree
 - Partially agree
 - Agree
 - Totally agree
- Q9) How adequate do you believe the payment you will receive for these activities is?
 - Not at all adequate
 - Very little adequate
 - Somewhat adequate
 - Fairly adequate
 - Moderately adequate
 - Quite adequate
 - Totally adequate
- Q10 DOUBLE) What effect did the payment of the bonus have on your performance in the second session?
 - I increased my performance because I felt my work was valued
 - I increased my performance because I felt grateful to the employer hiring for these activities
 - I decreased my performance because I felt I had done enough in the first session
 - I decreased my performance because I felt the employer was satisfied with my work
 - My performance was the same in the first session and in the second session
 - Other:
- Q10 CONTROL) How do you evaluate your performance in the second session?
 - I increased my performance because I felt my work was valued
 - I increased my performance because I felt grateful to the employer hiring for these activities
 - I decreased my performance because I felt I had done enough in the first session
 - I decreased my performance because I felt the employer was satisfied with my work
 - My performance was the same in the first session and in the second session
 - Other:
- Q11) Do you believe you will be offered to be hired again in the coming weeks within this same project?
 - It's certain that I will be offered another job opportunity

- It’s likely that I will be offered another job opportunity
- I wouldn’t know
- It’s unlikely that I will be offered another job opportunity
- It’s impossible that I will be offered another job opportunity, because the employer already clarified that the collaboration would be for one day only.
- Q12) Sex: M F
- Q13) Age: _____ years
- Q14) What is your marital status?
 - Married/Living with partner
 - Single
 - Separated/Divorced/Widow
- Q15) According to your utility bills, what is the tier of your current home or neighborhood?
 - | | | | | | |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
- Q16) What is your level of education?
 - None
 - Primary School
 - High school
 - Some university, but no graduation
 - Technical
 - University degree
 - Graduate degree
- Q17) What is your father’s level of education?
 - None
 - Primary School
 - High school
 - Some university, but no graduation
 - Technical
 - University degree
 - Graduate degree
 - No relationship with father
- Q18) What is your mother’s level of education?
 - None
 - Primary School
 - High school
 - Some university, but no graduation
 - Technical
 - University degree
 - Graduate degree
 - No relationship with mother

- Q19) What is your occupation?
 - Director or manager
 - Scientific or intellectual professional
 - Mid-level technician or professional
 - Administrative support staff
 - Service worker or salesperson in shops and markets
 - Farmer or skilled worker in agriculture, forestry, and fishing
 - Craftsman, operator, or artisan in mechanical arts and other trades
 - Operator of assembly plants and machines
 - Basic occupation
 - Military occupation
 - Unemployed, retired, housekeeper
 - Student
 - Other (specify):
- Q20) For each of the statements below, please indicate your level of agreement or disagreement.
 - a. The work organization was effective.

*	Strongly disagree	1	2	3	4	Strongly agree
---	-------------------	---	---	---	---	----------------
 - b. The payment I received was appropriate.

*	Strongly disagree	1	2	3	4	Strongly agree
---	-------------------	---	---	---	---	----------------
 - c. The treatment received by the two people hired was equitable.

*	Strongly disagree	1	2	3	4	Strongly agree
---	-------------------	---	---	---	---	----------------
 - d. The earnings I received in the second part of the workday were well deserved.

*	Strongly disagree	1	2	3	4	Strongly agree
---	-------------------	---	---	---	---	----------------
 - e. The earnings that the other person hired in the second part of the workday received were well deserved.

*	Strongly disagree	1	2	3	4	Strongly agree
---	-------------------	---	---	---	---	----------------