Behavioral Externalities of Process Automation

Ruth Beer
Zicklin School of Business, Baruch College, City University of New York, New York, NY 10010
ruth.beer@baruch.cuny.edu

Anyan Qi
Naveen Jindal School of Management, The University of Texas at Dallas, Richardson, TX 75080,
axq140430@utdallas.edu

Ignacio Rios
Naveen Jindal School of Management, The University of Texas at Dallas, Richardson, TX 75080,
iar200000@utdallas.edu

We study the behavioral effects of process automation on human workers interacting with automated tasks. We introduce a stylized model with two workers who complete their tasks sequentially, working toward a joint project to obtain a fixed payment plus a variable bonus that depends on how early the project is completed. We show that workers will complete their tasks as soon as possible if the early completion bonus is high enough. Following the behavioral operations literature, we hypothesize that workers will suboptimally delay project completion. In addition, we predict that automation will alleviate this problem by reducing the uncertainty in regard to task completion, leading to a higher project completion rate and worker productivity. To test these predictions, we conduct an experiment replicating the theoretical model, varying whether a worker collaborates with a coworker or robot. First, we find that workers largely deviate from the optimal policy, as they take longer than what the theory prescribes to complete their tasks or do not complete the project. Second, we show that process automation increases the project completion rate and reduces the project completion time, confirming the benefits of process automation. Interestingly, workers who collaborate with robots take *longer* to complete their tasks, contradicting our initial hypothesis that process automation has a positive effect on the productivity of human workers. An additional treatment shows that the reduced uncertainty derived from process automation cannot be leveraged to improve workers' productivity in the same way as reduced uncertainty in a human-human setting and that social preferences are an important driver of this result.

Key words: behavioral operations management, experiments, process automation, collaboration.

History: file version December 6, 2022

1. Introduction

The increasing interaction between humans and robots in the workplace has raised a call for more Operations Management research on the behavioral effects of process automation. The amount of

human-robot interactions has increased in recent years, aided by new technologies that allow for the development of "cobots" (or collaborative robots) that are capable of working safely alongside humans (Olsen and Tomlin 2020). For example, BMW reports that they are implementing robots that can work safely alongside humans on the production line (Knight 2014). Although, in earlier stages of process automation, robots were kept behind fences or in cages to avoid harming humans, new sensor technologies allow for safe close collaborative work between humans and robots. In this regard, Starbucks has recently announced that it has started to test robots that can feed ice into their equipment and can brew and pour coffee into individual cups, alongside humans (Russ 2022). Similarly, Chipotle is testing a new autonomous kitchen assistant to handle baking its tortilla chips, a core menu item (Rogers 2022). While these examples involve physical robots, virtual robots (or "bots") also are increasingly being used to automate time-consuming, repetitive tasks (which are usually tedious for human workers) through robotic process automation (RPA). For instance, DHL (2022) reports a successful RPA project, in which a bot extracts data from DHL's operating system and combines it with critical flight data to identify potential delays or other problems with the cargo. The RPA bot then generates a report that is passed on to human workers who decide the best course of action to address the issue. These examples of close interactions between human workers and robots or algorithms suggest that process automation may directly (and indirectly) affect human decisions and actions. Thus, the goal of this research is to gain a better understanding of the behavioral effects of process automation on human workers who interact with an automated task, as well as how to improve the collaboration between humans and robots/bots.

There is a growing literature on new trends in the labor market and the "future of work" with a focus on human-algorithm connections (Caro et al. 2022). This literature explores topics such as algorithm aversion and algorithm reliance (Christin 2017, Cao and Zhang 2020, Lin et al. 2022, Snyder et al. 2022), algorithm transparency (Diakopoulos 2020, Lebovitz et al. 2022) and the extent to which humans should be allowed to override algorithm recommendations (Ge et al. 2021, Ibrahim et al. 2021, Caro and Saez de Tejada Cuenca 2022). Most of this literature considers settings in which the human and the algorithm complete the same task, leading to the question of whether the human's or the algorithm's recommendation (or a combination of both) should be implemented. In this paper, we explore the behavioral effects of human-robot collaborations when they work on *different*, but interrelated, tasks, and we investigate two main research questions: (i) *What are the externalities of process automation on the performance of human workers who collaborate with a robot?* and (ii) *What behavioral factors affect human workers' response to collaborating with a robot?*

To study these questions, we consider a stylized collaborative setting with two sequential tasks, and compare the setting where each task is completed by a human worker to a process-automation setting where a human worker collaborates with a robot. In particular, we are interested in settings in which it is in the workers' best interest to work on the task (from a profit maximization standpoint), but they may be enticed to perform an alternative activity that delays the task completion. In such settings, process automation has the advantage of reducing the average processing time and the uncertainty about task completion, as the robot does not delay the completion of the task. Little is known, however, about the behavioral effects that process automation may have on the human workers who interact with the robot and the resulting overall productivity.[1] We conjecture that process automation will have a positive effect on the productivity of human workers and alleviate potential delays, as collaborating with a robot eliminates the strategic uncertainty of human-human interactions. We first build an analytical model and derive formal hypotheses. Then we test the hypotheses with a behavioral experiment.

To set a baseline, we first introduce a stylized theoretical model with a project that consists of two sequential tasks. One worker is assigned to complete the first task, and another worker is assigned to complete the second task, which cannot start until the first task is completed. Each task takes a certain number of periods to complete, and the project is completed when the second task is finished. If workers complete the project before a deadline, both workers earn the project completion payoff, which includes a fixed amount and a variable component that increases linearly in the number of periods before the deadline that the project was completed. This setup captures the collaborative nature of the work, as the two workers' payoffs are highly interrelated. To capture workers' tendency to delay their tasks, we assume that workers can take a (random) alternative option (high or low) in each period, which generates a private benefit to the worker but does not contribute to the project. Therefore, when a worker's task is active (i.e., the task has started but has not yet been completed), the worker chooses to work on the project or take the alternative option. When the task is not active, the worker earns the value of the alternative option. The alternative option captures other uses of the worker's time (e.g., working on another project or spending time on a recreational activity) and may generate an incentive to delay the task.

We derive the subgame-perfect equilibrium by backward induction and the optimal policy about whether to work on the project or take the alternative option from a stochastic dynamic program

[1] Process automation has other important potential benefits, including improved quality and reduced costs (Davenport and Ronanki 2018, Moffitt et al. 2018). In this paper, we focus on two main features of process automation: shorter processing time and reduced uncertainty. We leave the study of other aspects of process automation to future research.

for each worker. We show that, if the per-period early completion bonus is high enough, both workers complete their task as soon as possible, conditional on having enough time to do so. If the bonus is not high enough, workers can be better off by delaying their task completion when they face a high alternative option, provided that they still have enough time to complete the tasks after the delay.

To study the behavioral effects of process automation in this setting, we conduct an experiment and compare a human-human baseline where both tasks are performed by human workers with the human-robot treatment where the first task is performed by a human while the second task is performed by an automated robot. The automated robot is set to always complete the task in the minimum required number of periods (i.e., there is no delay). We carefully chose the values of the parameters so that it is optimal for human workers to finish the project as early as possible (and, therefore, the optimal policy of the second human worker should be the same as that of the automated robot), while setting the realized values of the alternative option high enough to avoid trivial cases in which the alternative options are not attractive to human workers. We hypothesize that, consistent with previous behavioral operations literature showing that deviations from the optimal policy are common in dynamic optimization settings (e.g., Palley and Kremer 2014, Long et al. 2016, Kagan et al. 2021), the subjects in our setting may be enticed to (suboptimally) take the alternative option and delay the task. Process automation should mitigate such delays due to the actions of the second worker and lead to a higher project completion rate and a shorter project completion time. Further, because the automated robot has no delay, a human worker who collaborates with the automated robot faces no uncertainty in regard to the duration of the second task. In contrast, in the human-human setting, a worker who performs the first task is exposed to strategic uncertainty. This suggests that process automation may have the additional benefit of improving the productivity of the human worker who collaborates with the automated robot, as it eliminates the worker's concern about the project's not being completed due to the coworker's delay.

In line with our hypothesis, the experimental results show significant deviations of human subjects from the optimal policy, i.e., the average task duration is significantly longer than what the normative theory prescribes, and the project completion rate is significantly below 100%. Process automation helps to mitigate these effects, leading to a higher project completion rate and shorter project duration. Interestingly, we find that process automation has a negative effect on the productivity of the human workers, who take longer to complete their tasks when the second task is automated compared to the case in which they collaborate with another worker. We

conduct an additional treatment to compare the effects of reduced uncertainty that results from process automation with the effects of reduced uncertainty in a human-human setting. We find that the behavioral response is context-specific: Reduced uncertainty about a later task duration in a human-human setting (that results from a coworker's commitment to finishing fast) has a more positive effect than does reduced uncertainty that results from process automation. Further analysis shows that social preferences are an important driver of the results, as pro-social subjects tend to be more productive when collaborating with a human coworker than with a robot.

The contribution of our paper is threefold. First, we introduce a framework to analyze the effects of process automation on human workers who interact with robots/bots. To the best of our knowledge, this paper is the first in the operations management literature to explore the behavioral externalities of process automation on the productivity of human workers who perform a collaborative task. Second, we conduct an experiment to empirically study the effects of process automation and analyze the behavioral drivers of the observed results. Finally, we provide concrete managerial insights derived from our experimental results that can be applied by companies when deciding whether to automate tasks, which tasks to automate, and how to design a collaborative work environment for the "future of work".

The remainder of this paper is organized as follows. We discuss the related literature in Section 2. In Section 3, we present the analytical model and derive normative predictions. In Section 4, we describe the experimental design and outline the main hypotheses. In Section 5, we report the experimental results, and in Section 6, we conclude.

2. Literature Review

Our paper lies at the intersection of three different streams of literature. First, our paper contributes to the growing literature exploring the human-algorithm connection. Second, our paper relates to the Behavioral Operations Management (BOM) literature on dynamic decision making. Finally, our paper is also broadly related to the BOM literature on collaborative projects.

Human-Algorithm Connection. The increased preponderance of algorithms for managerial decision-making has led to a growing literature on the human-algorithm connection. Several topics have emerged in this literature seeking to understand the coexistence of humans with algorithms and how to improve it (Caro et al. 2022). Among these topics, an important question is how to improve the adherence to algorithmic recommendations and overcome "algorithm aversion" (Dietvorst et al. 2015). This question has been studied in the context of retail markdowns recommendations (Caro and Saez de Tejada Cuenca 2022), algorithmic prescriptions for order packing

instructions of warehouse workers (Sun et al. 2022), the daily work of web journalists and legal professionals (Christin 2017), sales of an online education platform (Cao and Zhang 2020), algorithm use in type 1 diabetes self-management (Lin et al. 2022), and service workers making personalized recommendations (Snyder et al. 2022). A related stream of literature explores the hindering effects of algorithm opacity (the "black box issue") on the adherence to its recommendations (Diakopoulos 2020). Lebovitz et al. (2022) showed that opacity was detrimental for diagnostic radiologists using AI tools, while Siemsen and Aloysius (2020) discuss how managers' trust in systems that are not transparent can be an issue in supply chain decision-making. Several papers further explore whether (and how) humans should be allowed to override algorithmic recommendations. For example, Ibrahim et al. (2021) study how to use human judgement to determine the adjustments that should be made to an algorithmic forecast to account for information that the human has and the algorithm does not use. Dietvorst et al. (2018) find that humans show a preference for modifiable algorithms, while Ge et al. (2021) show that human-in-the-loop interferences of investors using robo-advisors often lead to inferior performance. With the exception of Bai et al. (2022), who study the effects of algorithmic task assignment on warehouse workers' performance, most of the literature on the human-algorithm connection focuses on settings where the human and the algorithm complete the same task. In these cases, a key question is how to reconcile the algorithmic recommendation with the human-judgement when they diverge. We contribute to this literature by exploring the behavioral effects of process automation in a collaborative environment, where a human and a robot work on *different*, but interrelated, tasks.

Behavioral Operations in Dynamic Decision Making. Among the variety of dynamic decision-making problems in the behavioral operations literature, our problem generally fits in the intertemporal time preferences family of problems, where agents face a trade-off between immediate and (potentially higher) delayed earnings. Although a large part of this literature focuses on savings and consumption decisions (see Duffy and Li (2019) for a review), there are applications in the BOM literature, such as revenue management, inventory management, and capacity decisions. For instance, Bearden et al. (2008) consider a seller with a fixed inventory who faces a sequence of buyers that arrive sequentially over time and find that revenue losses are driven mainly by agents being too demanding when holding higher levels of inventory and insufficiently demanding when holding low inventory levels. Kim et al. (2020) analyze a similar problem in hospital admissions and find that physicians underestimate the number of empty beds and that decision noise affects their behavior, leading them to over or under-utilize the system depending on the setting. Leider and Şahin (2014) study a setting where agents must allocate a fixed budget over a series of

alternatives that arrive over time and find that agents consistently exhibit an overuse bias, i.e., they overestimate the value of early options and overconsume their budget with low-value options. Palley and Kremer (2014) analyze a sequential search problem and show that limited information induces longer search and observe systematic deviations in terms of the timing to stop the search. More recently, Kagan et al. (2021) study dynamic decision-making in several well-known operations management problems and explore whether subjects are forward-looking optimizers in dynamic environments. The authors find that, depending on the features of the problem, agents may use static policies (for instance, in stopping problems) or simple forward-looking policies that depart from optimality (for instance, in technology adoption or capacity allocation problems). In contrast with this stream of literature, which typically considers a single decision-maker who makes dynamic decisions, we analyze the problem of multiple workers collaborating on sequential projects in which each worker makes dynamic decisions. Furthermore, we focus on analyzing the behavioral consequence of process automation, which has not been studied in the aforementioned literature.

Behavioral Operations in Collaborative Projects. Our work is also related to the behavioral operations literature on collaborative projects; see Grushka-Cockayne et al. (2018) for a review. Several papers employ analytic models incorporating behavioral biases of the decision makers or prescribing solutions to improve efficiency or mitigate delays in the context of project management and new product development, such as the investment and innovation sharing mechanisms (Bhaskaran and Krishnan 2009), the effect of cost salience (Wu et al. 2014), setting milestones and deadlines (Zhang 2016), time-based incentive contracts (Kwon et al. 2010, Rahmani et al. 2017, Dawande et al. 2019, Song et al. 2021, Chen et al. 2021, Hou et al. 2021), encouraging help across projects (Crama et al. 2019), and the framework with managerial mental accounting (Baucells et al. 2018). More recently, there are papers combining both analytical models and lab experiments to provide a closed-loop analysis on this subject, such as the acceptance and continuation decision for innovative projects (Wuttke et al. 2018), improving coordination among contractors with feedback (Shokoohyar et al. 2019), sequential abandonment for projects with uncertain value (Long et al. 2016), and the impact of communication between firms (Beer and Qi 2021). We contribute to this literature by studying the effects of process automation, which supposedly should mitigate delays, on workers' productivity in collaborative projects where agents are prone to delay the completion of their tasks.

3. Model

In the benchmark model, we consider a project that consists of two tasks to be performed sequentially by two workers respectively. Specifically, task 1 is performed by worker W_1 and must be

completed before worker W_2 can start working on task 2. Each worker W_i needs to spend τ_i periods of work to complete their task, and the deadline to complete the project is period T (with $T > \tau_1 + \tau_2$). In each period t that W_i's task is active, i.e., W_i can work on the task and has not completed it, the worker can choose to work on the task (which we denote by $x_{i,t} = 1$) or take an alternative option (which we denote by $x_{i,t} = 0$) that provides the worker a private value $a_{i,t}$ but does not contribute towards the completion of the task/project. The value of the alternative option is a random variable that takes value $a_{i,t} = h$ w.p. $p \in (0,1)$ and $a_{i,t} = l$ w.p. $1-p$ (with $h > l$), and this random variable is independent for each period and worker. Let $\mathbf{a}_t = (a_{1,t}, a_{2,t})$ be the vector of the random alternative options for W_1 and W_2 in period t, and $\mathbb{E}[a] = hp + l(1-p)$ denote the expected value of the alternative option in each period. The value of the alternative option is realized at the beginning of each period and observed by W_i before making their decision. In each period t that W_i's task is not active, the worker earns the realized alternative option.

Let $r_{i,t}$ be the number of periods that W_i has left to complete their task at the beginning of period t with $r_{i,0} = \tau_i$. Then, if the task is active at the beginning of period t, we have that $r_{i,t+1} = r_{i,t} - x_{i,t}$. That is, if W_i works on the project in period t, then the number of periods left to complete the task decreases by 1; otherwise, it does not change. Moreover, the project is completed in period t if $r_{2,t} = 1$ and $x_{2,t} = 1$. In that case, W_i earns a total profit of $\pi_i + \delta_i \cdot (T - t)$, where π_i is the value of the project for W_i and δ_i is an early termination bonus that W_i gets for each period that the project is completed ahead of the deadline. If the project is not completed by the deadline, both workers earn 0 from the project.

Let $\mathbf{r}_t = (r_{1,t}, r_{2,t})$ be the vector of remaining periods to complete tasks 1 and 2. Given the sequential nature of the project, we know that either (i) $r_{1,t} > 0$ and $r_{2,t} = \tau_2$, or (ii) $r_{1,t} = 0$ and $r_{2,t} \geq 0$ hold in each period. In the former case, we say that task 1 is active, and thus W_1 can work on it. In the latter case, $r_{1,t} = 0$ implies that W_1 has completed the first task, and thus W_2's task is active if $r_{2,t} > 0$. Finally, if $r_{1,t} = 0$ and $r_{2,t} = 0$, then both workers already completed their tasks and the project is completed. In what follows, we denote the set of feasible vectors $(r_{1,t}, r_{2,t})$ by

$$R = \{(r_1, r_2) \in \{0, \ldots, \tau_1\} \times \{0, \ldots, \tau_2\} : r_1 > 0 \text{ and } r_2 = \tau_2, \text{ or } r_1 = 0 \text{ and } r_2 \geq 0\}.$$

In the sequel, we drop the time index t of $\mathbf{r}_t$, $\mathbf{a}_t$, and $x_{i,t}$ when there is no ambiguity in doing so.

In Assumption 1 below, we assume that the value of the project π_i for W_i exceeds the expected value that the worker would have obtained had they chosen to take the alternative options for the duration of their task, and the alternative option is high in the current period. This assumption guarantees that workers are incentivized to complete their tasks if they have enough time.

ASSUMPTION 1. $\pi_i \geq h + \mathbb{E}[a](\tau_i - 1)$.

Let $V_{i,t} : R \times \{l,h\}^2 \to \mathbb{R}$ be a function such that $V_{i,t}(\mathbf{r},\mathbf{a})$ represents the optimal profit-to-go of W_i in period t when the number of periods to complete the tasks is $\mathbf{r} = (r_1, r_2) \in R$ and the alternative option is realized as $\mathbf{a} = (a_1, a_2) \in \{l,h\}^2$. For $i = 1,2$, and $t = 1,2,...,T$, we have

$$\begin{aligned} V_{i,t}(\mathbf{r},\mathbf{a}) &= \max_{\substack{x_i \in \{0,1\} \\ x_1 \leq \mathbb{1}_{\{r_1>0\}} \\ x_2 \leq \mathbb{1}_{\{r_1=0\}}}} \left\{ (1-x_i)\cdot a_i + \mathbb{1}_{\{r_2=0\}} \cdot \delta_i + \mathbb{E}_{\mathbf{a}_{t+1}}\left[V_{i,t+1}\left((r_1 - x_1, r_2 - x_2), \mathbf{a}_{t+1}\right)\right] \right\} \\ V_{i,T+1}(\mathbf{r},\mathbf{a}) &= \pi_i \cdot \mathbb{1}_{\{r_2=0\}}. \end{aligned} \tag{1}$$

When task 1 (resp., task 2) is active, we have $\mathbb{1}_{\{r_1>0\}} = 1$ (resp., $\mathbb{1}_{\{r_1=0\}} = 1$). The constraints $x_1 \leq \mathbb{1}_{\{r_1>0\}}$ and $x_2 \leq \mathbb{1}_{\{r_1=0\}}$ imply that W_i can work on task i (i.e., $x_i = 1$) only when their task is active. It is easy to see that since tasks 1 and 2 cannot be active in the same period, the two workers will not work on their tasks in the same period, i.e., $x_1 \cdot x_2 = 0$.

We derive the subgame-perfect equilibrium of the game by backward induction. In Proposition 1 we characterize the optimal strategy for W_2. Similarly, in Proposition 2 we characterize the optimal strategy for W_1.

PROPOSITION 1 **(Optimal Strategy for W_2)**. *If $\delta_2 \geq h - \mathbb{E}[a]$, then the optimal strategy for worker W_2 is:*

$$x^*_{2,t}(\boldsymbol{r},\boldsymbol{a}) = \begin{cases} 1 & \text{if} \quad r_1 = 0 \text{ and } 0 < r_2 \leq T - t + 1; \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

By contrast, if $\delta_2 < h - \mathbb{E}[a]$, then the optimal strategy for worker W_2 is:

$$x^*_{2,t}(\boldsymbol{r},\boldsymbol{a}) = \begin{cases} 1 & \text{if} \quad r_1 = 0,\ 0 < r_2 \leq T - t + 1, \text{ and either } a_2 = l \text{ or } r_2 = T - t + 1; \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

PROPOSITION 2 **(Optimal Strategy for W_1)**. *Consider W_2 follows the optimal strategy characterized in Proposition 1. If $\delta_1 \geq h - \mathbb{E}[a]$, then the optimal strategy for worker W_1 is:*

$$x^*_{1,t}(\boldsymbol{r},\boldsymbol{a}) = \begin{cases} 1 & \text{if} \quad 0 < r_1 \leq T - \tau_2 - t + 1; \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

By contrast, if $\delta_1 < h - \mathbb{E}[a]$, then the optimal strategy for worker W_1 is:

$$x^*_{1,t}(\boldsymbol{r},\boldsymbol{a}) = \begin{cases} 1 & \text{if} \quad 0 < r_1 \leq T - \tau_2 - t + 1, \text{ and either } a_1 = l \text{ or } r_1 = T - \tau_2 - t + 1; \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Propositions 1 and 2 imply that, if the early completion bonus (δ_i) exceeds the difference between the high alternative option and the expected value of the alternative option, both workers complete their task as soon as possible conditional on having enough time to complete it. If the early completion bonus is less than the difference between the high alternative option and the expected

value of the alternative option, both workers still have incentives to complete the project (as Assumption 1 holds), but workers are better off by delaying the task/project completion if they face a high alternative option and still have enough time to complete the task/project after the delay. Hence, the workers will work on the project if they encounter a low alternative option or are time-constrained to complete the project, i.e., if not working on the project rules out the chances to complete it. As a result, the policies for W_1 and W_2 are almost identical and only differ on the *effective* deadline, which is $T - \tau_2$ for W_1 and T for W_2.

4. Experimental Design

We conducted an incentivized behavioral experiment to test our theoretical predictions. The experiment consists of ten rounds of a main game capturing the setting described in the analytical model (Section 3).[2] In addition, the subjects completed an exit survey about their decisions in the main game[3] as well as their demographical information, and two short additional tasks (to elicit measures of pro-sociality and cognitive ability). More specifically, the main game consists of completing a project with two sequential tasks. To contextualize the setting, subjects were instructed that the goal of the project was launching a new product. Our experimental design consists of a *Human-Human* baseline (**HH**), where both tasks are completed by human subjects, and a *Human-Robot* treatment (**HR**), where the first task is completed by a human subject and the second task is completed by an automated robot. The experiment was conducted following a between subject design; that is, subjects were only exposed to one condition. We next explain the HH and HR conditions in detail.

In the HH condition, at the beginning of each session, subjects are randomly assigned the role of Worker 1 or Worker 2, and they keep their role for the ten rounds of the game. Subjects in the role of Worker 1 and subjects in the role of Worker 2 are randomly and anonymously matched in pairs at the beginning of each round. To complete the project, workers must sequentially complete their respective tasks before the deadline. Specifically, Worker 1 performs task 1 and Worker 2 performs task 2, where each task takes three (not necessarily consecutive) periods. Given the sequential nature of the tasks, Worker 2 cannot start working on her task until Worker 1 completes task 1. The deadline to complete the project is twelve periods.

While their task is active, workers can choose in each period whether to work on the project or take an alternative option. The value of the alternative option for each worker is randomly drawn

[2] The detailed instructions and screenshots are presented in Appendix B.1.

[3] We analyze subjects' comments in the exit survey to get further insights on the behavioral drivers in Appendix C.1.

in each period, and it takes a value of 40 points (high) or 10 points (low), both with equal chance. Workers learn the realized value of the alternative option at the beginning of each period. When their task is not active and in the periods between the project completion and the deadline, both workers earn the value of the alternative option. Finally, if the project is completed by the deadline, each worker earns 100 points from the value of the project (otherwise, the project value is zero points). In addition, if the workers complete the project before the deadline, each of them earns a linear early completion bonus equal to $\delta = 30$ points times the number of early completion periods.

The values of the parameters are set so that (i) Assumption 1 in the model (Section 3) holds (therefore, completing the project is always optimal), and (ii) the early completion bonus exceeds the difference between a high alternative option and its expected value (therefore, both workers have an incentive to complete their task as soon as possible conditional on having enough time to complete it).

In the HR treatment, all the subjects are assigned the role of Worker 1 and are instructed that they are collaborating with an automated robot to complete a project. Subjects know that the automated robot will take exactly three periods to complete task 2, provided there is enough time to complete the project before the deadline (that is, as long as Worker 1 finishes task 1 by period 9). If task 1 is not completed by period 9, the project will not be completed. The values of the parameters and the payoff scheme are identical to the HH treatment.

Beliefs Elicitation. In the HH treatment, we elicit at the beginning of each round the workers' beliefs about the number of periods their coworkers will take to complete their tasks. The belief elicitation is incentivized with an additional payoff of 50 points if a subject correctly guessed the coworker's task duration.

Additional Tasks. After playing the ten rounds of the main game, subjects participated in two additional tasks. The *modified dictator game* (Andreoni and Vesterlund 2001, Leider and Lovejoy 2015) was conducted to elicit subjects' social preferences. In this game, subjects make five allocation decisions of 50 tokens between themselves and another participant, where the decisions differ in how much a token is worth to the allocator and to the recipient (ranging from \$0.05 to \$0.15). Subjects are then randomly assigned the role of allocator or recipient, and are randomly paired with a subject in the opposite role. The payoffs are determined for each pair according to one of the allocator's five decisions, randomly selected.

The next task consisted of answering four questions from the traditional *CRT game* (Frederick 2005) (or similar), aimed at eliciting a measure of the subjects' cognitive ability. The task

was incentivized with a payoff of \$0.5 for each correct answer. The questions are included in the Appendix B.2.[4]

4.1. Hypotheses

As previously discussed, the values of the parameters used in the experiment are set so that it is optimal for the subjects to (i) complete the project and (ii) work on the project whenever their task is active. Therefore, our normative theory prescribes that each task is completed in three periods, and the project is completed in six periods. Nevertheless, since we are interested in a case where the workers tend to delay the task, we also carefully selected the value of the early completion bonus to be not too high (relative to the alternative option) so that the subjects may be enticed to take the alternative option. Previous literature has shown that deviations from the optimal policy are common in dynamic optimization settings (Kagan et al. 2021). In particular, subjects are biased in deciding whether to continue or abandon a project (Long et al. 2016) and tend to stop too early/late in solving optimal stopping problems (Palley and Kremer 2014). In our baseline HH setting, the incentives to delay the task may be further exacerbated by the social interactions between the two workers. On the one hand, Worker 1 faces strategic uncertainty about whether Worker 2 will follow the optimal policy. As a result, Worker 1 may be incentivized to take the alternative option out of concerns that Worker 2's delay may result in the project not being completed. Worker 2 does not face strategic uncertainty (since, by the time Worker 2 starts working on the project, the duration of task 1 is already determined). However, Worker 2 may still be enticed to delay her task, particularly if she observes that Worker 1 already delayed the completion of the first task. Therefore, we expect that both workers are likely to deviate from the optimal policy.

Hypothesis 1 (Deviations from Optimal Policy.)

1. *The project completion rate is lower than what the optimal policy prescribes (100%).*
2. *The project completion time is longer than what the optimal policy prescribes (six periods).*

To test the effects of process automation in a setting where the workers have a tendency to delay the task, we designed a HR treatment, where the second task is performed by an automated robot that takes exactly three periods. Since process automation eliminates the potential delay in the completion of task 2, we expect that the HR treatment will result in a higher project completion

[4] The subjects also participated in a *trust game* (Berg et al. 1995) and a *lottery game* (Dohmen and Falk 2011), which allowed us to elicit additional measures of trustworthiness and risk aversion. The discussion of these behavioral factors is omitted as they have no significant predictive power on the main variables of our experiment.

rate and a shorter project completion time than the HH baseline. In addition, we conjecture that, by eliminating the uncertainty in the duration and completion of the second task, process automation will have a positive effect on Worker 1, who is no longer concerned about the project not being completed before the deadline. As a result, we hypothesize that the HR treatment will result in a shorter task 1 completion time than the HH baseline—that is, there is a *positive externality* of process automation on the productivity of Worker 1.

Hypothesis 2 (Effects of Process Automation.) *Process automation leads to:*

1. *Higher project completion rate.*
2. *Shorter project completion time.*
3. *Shorter task 1 completion time (higher productivity of Worker 1).*

5. Experimental Results

The experiment was conducted virtually between April and September 2022. The protocol was adapted from Zhao et al. (2020) to approximate a typical lab setting (Li et al. 2021) by combining an online experimental platform (z-Tree unleashed, Duch et al. 2020) with a parallel web-conferencing session (implemented on Zoom). The subjects were undergraduate and graduate students at a public university in the South of the US recruited through the online recruitment system SONA. Sessions lasted for approximately 75 minutes, and subjects' average payment was $18 (including a $5 show-up fee). Overall, 86 subjects participated in the experiment; 54 participated in the HH condition (27 in the role of Worker 1 and 27 in the role of Worker 2), and 32 participated in the HR condition (all of them in the role of Worker 1). No subject participated in more than one session.

5.1. General Results

Table 1 summarizes the main findings of the HH and HR conditions. The table reports the completion frequency (columns 1 to 3) and duration (columns 4 to 6) of each task and the overall project. The reported values correspond to subject-level averages, that is, one observation is a subject's average across the ten rounds played. More specifically, we compute the frequency with which each worker completes their task (number of times they complete their task divided by ten) and the average number of periods they take to complete their task conditional on completing the project. As a result, we obtain a dataset with one observation per subject. Note that the project completion frequency coincides with the completion frequency of the second task, and that the average project duration is the sum of the durations of the first and second tasks.

We analyze these results in two parts: in Section 5.1.1 we test Hypothesis 1 and analyze the deviations from the optimal policy, and in Section 5.1.2 we test Hypothesis 2 and analyze the effects of process automation.

Table (1) Summary Statistics by Treatment and Round

		Completion Frequency			Duration		
Treatment	N	Task 1	Task 2	Project	Task 1	Task 2	Project
HH	27	0.989	0.870	0.870	5.051	4.464	9.515
HR	32	0.997	0.984	0.984	5.346	3.000	8.346

Notes: Subject-level averages reported. "Duration" in columns $4-6$ is conditional on project completion.

5.1.1. Deviations from Optimal Policy. Hypothesis 1 prescribes that subjects deviate from the optimal policy. More specifically, Hypothesis 1.1 predicts that the project completion rate is *lower* than the 100% rate that the optimal policy prescribes, and Hypothesis 1.2 predicts that the project completion time is *longer* than the six periods that the optimal policy prescribes.

First, we find that the experimental data supports Hypothesis 1.1. We observe that only 87% of the projects are completed in the HH condition. This frequency is significantly lower than 100% (t-test $p < 0.001$). In the HR treatment, 98% of the projects are completed, also significantly lower than 100% (t-test $p = 0.023$). These results indicate that subjects suboptimally choose not to complete the project in several opportunities.

We now focus on the observations where the project was completed and analyze the average duration of each task and the overall project. The results are presented in the last three columns in Table 1. We observe significant departures from the optimal policy under both conditions. In the HH condition, task 1 takes on average 5.051 periods and task 2 takes on average 4.464 periods, both significantly longer than 3 periods (both t-test $p < 0.001$). In the HR condition, task 1 takes on average 5.346 periods, also significantly longer than 3 periods ($p < 0.001$). Recall that in the HR condition, the second task is automated and is programmed to have a duration of three periods as long as it can be completed before the project deadline. As a result of the delays in the tasks completed by human subjects, the overall project duration is also longer than six periods under both the HH and HR conditions (9.515 under HH and 8.346 under HR, both $p < 0.001$). These results confirm that Hypothesis 1.2 is supported by the data.

Finally, in the HH condition, we observe that Worker 1 takes significantly longer than Worker 2 to complete their task (5.051 vs. 4.464, Wilcoxon signed-rank test $p = 0.020$). While we did not derive a formal hypothesis to compare Worker 1's delay and Worker 2's delay in completing their respective tasks, we anticipated that Worker 1 is subject to strategic uncertainty while Worker 2 is not. As a result, Worker 1 may have concerns about her efforts to complete the task on time being "wasted" if Worker 2 does not complete task 2 before the deadline. The experimental results are consistent with this conjecture. Also, note that the fact that Worker 1's delay is more salient than

Worker 2's delay supports our decision to automate the second task in the HR treatment, as we can expect the impact of process automation to be more salient on Worker 1 (who, *a priori*, has a higher tendency to deviate from optimality).

In line with the previous result, we also observed that the majority of the cases where the project was not completed in the HH treatment were attributed to Worker 1's delay. More specifically, in 34% of the cases where the project was not completed, Worker 1 took longer than 9 periods to complete task 1 and, therefore, Worker 2 did not have enough time to complete task 2. The reminder of the cases (where Worker 2 could have completed task 2 but chose not to do it) are also, for the most part, attributed to Worker 1 taking too long. In fact, in 59% of the cases Worker 1 took between 7 and 9 periods, while only in 41% of the cases Worker 1 took 6 periods or less.[5] Overall, these results further confirm that Worker 1's delay is more salient than Worker 2's delay, and support our decision to analyze the effect of process automation on Worker 1.

5.1.2. Effects of Process Automation. Hypothesis 2 predicts that process automation will result in (i) a higher project completion rate (Hypothesis 2.1), and (ii) a shorter project completion time (Hypothesis 2.2). To test these predictions, we compare the project completion rate and project completion time in the HH and HR conditions. Since in the HR treatment the second task is automated (and, therefore, there will be no delays in the completion of task 2), we expect that the project completion frequency will be higher in the HR treatment than the HH treatment, and the project duration will be shorter in the HR treatment than the HH treatment.

The experimental results support Hypotheses 2.1 and 2.2. We observe that the project completion rate is significantly higher in the HR treatment than in the HH baseline condition (98% vs. 87%, Wilcoxon rank-sum test $p < 0.001$). Second, we observe that the project duration is shorter in the HR treatment than the HH baseline (8.346 vs. 9.515, $p < 0.001$). These results confirm the benefits of process automation, leading to higher probability of project completion and a shorter project duration.

We test these results more formally with regression analysis in Table 2. The table pools data from subjects in the role of Worker 1 in the HH and HR treatments and reports the results of panel linear regressions with subject random effects and standard errors clustered at the session level. Column (1) presents a regression of a variable that takes value one if the project was completed and zero otherwise on a dummy variable for the HR treatment, while column (2) presents a regression of the project duration on the HR treatment dummy. All the regressions control for round, cognitive

[5] We take as a reference a "fair" task 1 duration of six periods, which allocates equal amount of time for each worker.

ability, and demographics.[6] First, in column (1) we observe that the HR treatment leads to a higher project completion rate than the HH baseline ($\beta = 0.103$, $p = 0.006$), consistent with Hypothesis 2.1. Second, in column (2) we find that the HR treatment results in a shorter project duration than the HH baseline ($\beta = -1.179$, $p = 0.003$), confirming Hypothesis 2.2.

Table (2) Regression results: HH vs. HR

	Project Completion	Project Duration	Task 1 Duration
	(1)	(2)	(3)
HR treatment	0.103***	-1.179***	0.300**
	(0.037)	(0.260)	(0.128)
Constant	0.874***	9.494***	5.027***
	(0.023)	(0.461)	(0.486)
CRT	Yes	Yes	Yes
Round	Yes	Yes	Yes
Demographics	Yes	Yes	Yes
Observations	590	590	550
Subjects	59	59	59

Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. The table pools data from subjects playing the role of Worker 1 in the HH and HR treatments. Significance reported: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

Finally, Hypothesis 2.3 predicts that process automation has a positive externality on the productivity of Worker 1. Column (3) in Table 2 presents a regression of the task 1 duration on the HR treatment dummy. We observe that, contrary to the prediction, the HR treatment leads to a *longer* task 1 completion time than the HH baseline ($\beta = 0.300$, $p = 0.019$). This result suggests a *negative* externality of process automation on the productivity of Worker 1 (Hypothesis 2.3 is not supported).

In summary, we observe significant deviations from the optimal policy. The project completion rate is lower than 100% under both the HH and HR conditions and workers take significantly longer than six periods to complete their tasks. Hence, we conclude that Hypotheses 1.1 and 1.2 are supported by the data. Moreover, we confirm the benefits of process automation in leading to a higher project completion rate and reducing the time it takes to complete the project, in support of Hypotheses 2.1 and 2.2. Nevertheless, we find that automating the second task has a *negative* externality on the productivity of Worker 1. Rather than incentivizing Worker 1 to behave closer to the optimal strategy, we observe that eliminating the uncertainty faced by Worker 1 further increases Worker 1's delay.

[6] We create a measure of subjects' cognitive ability by taking the number of correct answers in the CRT game as a continuous variable. We confirm that the results are similar if we consider the answers as a categorical variable. The demographics controls include gender, major (economics, business, and other) and age group (< 22 and ≥ 22).

5.2. Understanding Drivers of Behavior.

While we conjectured that the reduced uncertainty of the HR treatment would have a positive externality on the productivity of Worker 1 (since, when task 2 is automated, Worker 1 is not concerned that Worker 2 may delay the project to the extent that the project would not be completed), the experimental results show that task 1's duration is *longer* in the HR treatment than in the HH baseline. In this section, we explore more in depth the behavioral effects of reduced uncertainty on Worker 1.

5.2.1. Uncertainty about Completion. One potential explanation for why Worker 1 takes longer to complete task 1 when she has more certainty that the project will be completed is that, since workers know that the automated robot will complete task 2 with certainty (provided enough time to do so), they do not need to provide extra "wiggle room" for the completion of task 2 and, thus, they can spend some extra time themselves in completing task 1. A follow-up question is whether reduced uncertainty will *always* make Worker 1 work more slowly. More specifically, we are interested in exploring the following question: *Does reduced uncertainty about the task 2 completion lead to longer task 1 duration when Worker 1 collaborates with another human?*

To answer this question, we conducted a new *Human-Human with Commitment* treatment (**HH-C**). The HH-C treatment extends the HH baseline by allowing both workers to make a commitment to completing their tasks in a certain number of periods. The commitments are not binding (i.e., workers are not forced to obey their commitment), and are made at the beginning of each round and are immediately announced to the coworker (before the project starts). As a reminder, the commitments are also displayed on the subjects' screens throughout the round. All the other aspects of the HH-C treatment are identical to the HH baseline.

The goal of the HH-C treatment is to analyze the case where Worker 2 commits to a fast completion time thus reducing Worker 1's uncertainty about the project completion, resembling Worker 1's incentives in the HR treatment. Therefore, our focus is on the commitment of Worker 2 and on Worker 1's behavioral response to Worker 2's commitment to finishing fast.[7] In order to ensure that the HH-C treatment indeed leads to reduced uncertainty for Worker 1, we need to confirm that it satisfies three conditions: (i) First, there is heterogeneity in Worker 2's commitments, so that we can evaluate Worker 1's response to a commitment to finishing *fast*. (ii) Second, provided that some Workers 2 commit to finishing fast, those who commit to finishing fast actually do so. (iii) Finally, Worker 1 believes the commitment of Worker 2. If the three conditions are met, suggesting

[7] While our focus is on the commitment of Worker 2, we also elicited the commitment of Worker 1 to avoid any unwanted asymmetries in the experimental design.

that Worker 2's commitment to finishing fast reduces Worker 1's uncertainty, we would expect to find a similar behavior between the HR and HH-C treatments, as formalized in Hypothesis 3.

Hypothesis 3 (Reduced Uncertainty under HH-C vs. HR.) *The HH-C treatment with a commitment to finishing fast by Worker 2 results in equal task 1 duration as the HR treatment.*

Results. The HH-C treatment was conducted with the same subject pool and protocol as the HH and HR treatments during the same time periods. A total of 52 subjects participated in this treatment, 26 in the role of Worker 1 and 26 in the role of Worker 2. The duration of each session and the average payment were similar to the HH and HR conditions. Overall, we observe that the project completion rate and project duration of the HH-C treatment are not significantly different from the HH baseline (project completion rate: 87% in HH vs. 84% in HH-C, $p = 0.496$; project duration: 9.530 in HH vs. 9.166 in HH-C, $p = 0.226$). We formally test for differences between the HH and HH-C treatments using a regression analysis, and the results in Table 9 (in Appendix C.3) confirm that there are no significant differences across the two treatments in terms of project completion rate and the time of completion. Moreover, the results in this table show that there are no significant differences across the two treatments in the durations of task 1 and task 2. These observations suggest that the results for the human-human settings are robust, and that the HH-C treatment provides a good proxy for the workers' behavior in the HH treatment.

Next, we test whether Worker 2's commitment to finishing fast contributes to reducing Worker 1's uncertainty about the project completion. We do so by testing whether the following conditions are met: (i) there is heterogeneity in Worker 2's commitments, (ii) Workers 2 who commit to finishing their task fast actually do so, and (iii) Worker 1 believes the commitment of Worker 2.

First we observe that there is heterogeneity in Worker 2's commitment (mean = 5.376, stdev = 1.868, see Figure 5 in Appendix C.3 for the full distribution). Therefore, we are able to analyze separately the cases with fast and slow commitments. To do so, we define that Worker 2 made a commitment to finishing "fast" if she committed to finishing in less than six periods (55^{th} percentile of the distribution) and "slow" if she committed to finishing in six periods or more. We choose a cutoff of six periods because Worker 1 rarely takes longer than six periods to complete task 1. Therefore, a commitment by Worker 2 to finishing within six periods would eliminate the uncertainty about the project being completed before the deadline.[8] For the purpose of this analysis, we only focus on the data corresponding to cases where Worker 2 chose to commit (14 out of 260 observations where dropped because Worker 2 chose not to commit).

[8] The results remain qualitatively the similar if we use instead cutoff values 3, 4, or 5.

Table (3) Effect of Worker 2's Commitment on Worker 1's Belief: HH-C

	Worker 1's Belief	
	(1)	(2)
Worker 2's Commitment	0.376***	-
	(0.109)	-
Worker 2's Commitment to Finishing Fast	-	-1.169***
	-	(0.391)
Constant	5.068***	7.779***
	(1.282)	(0.498)
CRT	Yes	Yes
Round	Yes	Yes
Demographics	Yes	Yes
Observations	205	205
Subjects	26	26

Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. The table considers data from subjects playing the role of Worker 1 in the HH-C treatment when the project was completed and Worker 2 made a commitment. Significance reported: $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

Second, we observe that when Workers 2 commit to finishing "fast" (in less than six periods), they usually do so. Specifically, we find that 63% of the Workers 2 who committed to finishing task 2 fast finished their task by the committed time, 87% finished within the committed number of periods plus one, and 97% finished in six periods or less. This results indicate that commitments to finishing fast by Worker 2 are, for the most part, honored.

Finally, we explore whether Worker 1 believes Worker 2's commitment to finishing fast. In Table 3 we analyze the effect of Worker 2's commitment on Worker 1's beliefs. We elicit the workers' beliefs about the number of periods that their coworkers will take to complete their task at the beginning of each round, right after they observe the commitments made by their coworkers.[9] Column (1) considers the commitment made by Worker 2 as a continuous variable (ranging from 3 to 9), while column (2) considers a dummy variable equal to 1 if the commitment made by Worker 2 was fast (i.e., less than six periods) and zero otherwise. Both regressions show consistent results: the first column shows that Worker 1's belief is positively correlated with Worker 2's commitment ($\beta = 0.376$, $p = 0.001$), while the second column shows that Worker 1's belief is significantly lower when Worker 2's commitment is "fast" than when the commitment is "slow" ($\beta = -1.169$, $p = 0.003$). Taken together, these results suggest that Worker 1's uncertainty about the project completion is reduced when Worker 2 commits to finishing fast.

Having confirmed that in the HH-C treatment a commitment to finishing fast by Worker 2 leads

[9] The belief elicitation method is the same as that used in the HH baseline. That is, the belief elicitation is incentivized with an additional payoff of 50 points if a subject guessed the coworker's task duration correctly.

to a reduced uncertainty by Worker 1, we next test Hypothesis 3. More specifically, we explore whether the HH-C treatment with a commitment to finishing fast by Worker 2 results in an equal task 1 duration as the HR treatment. We observe that, contrary to the prediction of Hypothesis 3, the task 1 duration is significantly shorter when Worker 2 commits to finishing fast than in the HR treatment (4.817 vs. 5.346, Wilcoxon rank-sum test $p < 0.001$). Furthermore, we observe that the task 1 duration is significantly shorter when Worker 2's commitment is fast vs. slow (4.817 vs. 5.188, $p = 0.036$), and that a slow commitment results in a task 1 duration that is not significantly different from the HR treatment ($p = 0.464$). We test these results more formally in Table 4. The table pools data from the HH-C and HR treatment and reports the results of panel linear regressions of the task 1 duration on dummy variables for a fast and slow commitment in the HH-C treatment.

Table (4) Regression results: HH-C vs. HR

	Task 1 Duration
Worker 2's Commitment to Finishing Fast	-0.594***
	(0.144)
Worker 2's Commitment to Finishing Slow	-0.221
	(0.333)
Constant	4.911***
	(0.675)
CRT	Yes
Round	Yes
Demographics	Yes
Observations	520
Subjects	58

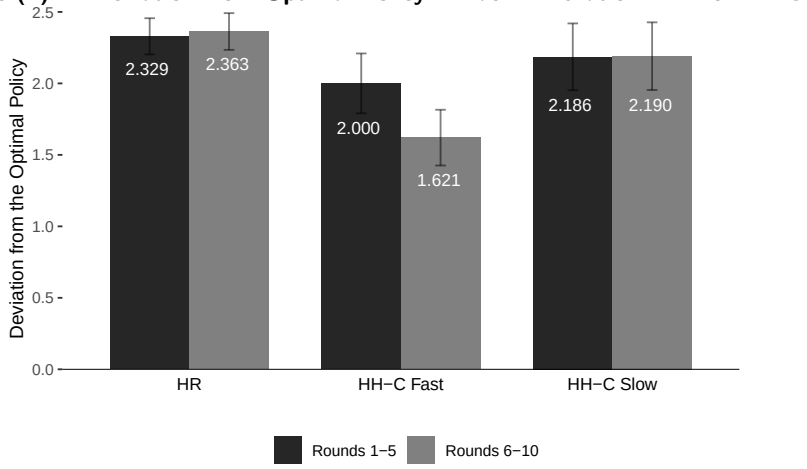
Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. The table considers data from subjects playing the role of Worker 1 in the HR and HH-C treatments when the project was completed and Worker 2 made a commitment. Significance reported: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

On the one hand, we confirm that Workers 1 who observe a fast commitment by Worker 2 take significantly less time to complete task 1 compared to the HR treatment. On the other hand, we observe that there are no significant differences in the task duration of Workers 1 who observe a slow commitment and those in the HR treatment. Taken together, these results suggest that a reduced uncertainty about the task 2 duration has a different impact on Worker 1 depending on how the reduced uncertainty originated: when it results from a coworker's commitment to finishing fast, the effect of reduced uncertainty on Worker 1's productivity is significantly more positive than when it results from process automation.

To explore the result more in depth, in Figure 1 we plot the deviations from the optimal policy in the duration of task 1 for the HR treatment and the HH-C treatment, separating the cases where

Worker 2 committed to finishing fast (in less than six periods) and slow (in six periods or more). To identify any potential learning effects, the figure plots separate rounds 1 to 5 and 6 to 10 of play in a session. We observe that in the HR treatment, the duration of task 1 is consistent as rounds in a session elapse. In the HH-C treatment, the positive effect of a fast commitment by Worker 2 on the duration of task 1 increases with the rounds of play. These results suggest that the subjects' learning strengthens the positive effect of Worker 2's fast commitment on Worker 1, while there is no significant effect of learning by Worker 1 when task 2 is automated.

Figure (1) Deviation from Optimal Policy in Task 1 Duration: HR vs. HH-C



Summary. We observed that while in human-human interactions reduced uncertainty about task 2 completion makes Worker 1 work faster, reduced uncertainty resulting from process automation makes Worker 1 work more slowly. Furthermore, we observed that the differences between these two cases are reinforced with learning. We conjecture that the difference between Worker 1's response to reduced uncertainty in the human-human case and the human-robot case may be explained by social preferences. In particular, if Worker 1 cares about the well-being of Worker 2, Worker 1 may be incentivized to work faster so that: (1) Worker 2 has more time to complete her task (this behavior would be consistent with preferences for *fairness*), (2) both workers earn the early completion bonus (consistent with preferences for *surplus maximization*). In the next section we explore the role of social preferences as a driver of the Worker 1's behavior.

5.2.2. Social Preferences. A primary feature of the HR treatment is that Worker 1 does not collaborate with another human worker. Therefore, if Worker 1 delays the completion of her task, she is not inflicting any penalty on a coworker. On the contrary, delaying the completion of task 1 in both the HH and HH-C treatments has a negative effect on a human collaborator. This

suggests that the longer duration of task 1 in the HR treatment relative to the HH and HH-C treatments may be explained by social preferences. In particular, two distinctive forms of social preferences may arise when task 2 is performed by a human worker. First, Worker 1 may have preferences for *fairness* (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) and may want to split the total number of periods available to complete the project in an equitable manner with Worker 2. Preferences for fairness have been identified in the BOM literature in settings such as supply chain contracting (Cui et al. 2007, Katok and Pavlov 2013, Ho et al. 2014, Kalkancı et al. 2014), retailer inventory-sharing practices (Davis et al. 2022), and procurement decision-making (Beer et al. 2021). Second, Worker 1 may have concerns for *surplus maximization* (Andreoni and Miller 2002) and derive utility from the combined payoff of both workers. As a result, Worker 1 would have an additional incentive to accelerate the task to earn a higher early completion bonus (which benefits both workers) over delaying the task to take a high alternative option payoff (which only benefits herself). Preferences for surplus maximization have also been studied in a supply chain context in the BOM literature (Beer et al. 2018).

To study whether social preferences can explain behavior in this setting, we classify subjects according to their social type based on the modified dictator game performed after the main game. More specifically, we classify subjects as (i) "selfish", if they show a tendency to keep all the tokens (36% of the subjects); (ii) "fairness maximizers", if they show a tendency to equalize payoffs (50%); and (iii) "surplus maximizers", if they tend to maximize the sum of their own and the other subject's payoffs (14%).[10] Since preferences for fairness maximization and surplus maximization would lead to the same behavior by Worker 1 in our setting (shorter task 1 completion time), we pool these two types and denote them more generally as "pro-social".

We hypothesize that the shorter task 1 duration in the HH and HH-C conditions (where task 2 is performed by a human) relative to the HR condition (where task 2 is automated) is due to Worker 1's social preferences. More specifically, we hypothesize that it is the prosocial subjects that drive the differences across treatments, as formalized in Hypothesis 4.

Hypothesis 4 (Social Preferences.) *The average task 1 duration of a pro-social Worker 1 is shorter in the HH and HH-C conditions than in the HR condition, while the average task 1 duration of a selfish Worker 1 in the HH and HH-C conditions is not different from the HR condition.*

[10] We refer the reader to Appendix C.2 for more details about this analysis.

Results. We first conduct a preliminary analysis to test whether the modified dictator game generated valid measures of subjects' social preferences (i.e., "selfish" or "pro-social" types) to explain the behavior in the main game. On the one hand, since Worker 1 in the HR treatment does not collaborate with a human worker, we expect no differences in the duration of task 1 between selfish and pro-social workers. This result is confirmed in Table 5, which shows that the average duration of task 1 in the HR treatment are 5.266 and 5.401 for selfish and pro-social workers, respectively (Wilcoxon rank-sum test $p = 0.581$) On the other hand, we expect that pro-social workers take less time to complete task 1 compared to selfish workers in the HH and HH-C treatments. From Table 5 we observe that the difference between selfish and pro-social subjects is directionally larger in the HH and HH-C conditions than in the HR condition (HH: 5.406 for selfish vs. 4.918 for prosocial, $p = 0.107$; HH-C: 5.312 for selfish vs. 4.824 for prosocial, $p = 0.012$).

Table (5) Time to complete task 1 by treatment and social type: HR vs. HH vs. HH-C

		% of Workers 1		Task 1 Duration	
	Treatment	Selfish/Pro-social	Obs.	Mean	SE
Selfish	HR	41%	128	5.266	0.136
	HH	26%	64	5.406	0.233
	HH-C	35%	77	5.312	0.178
Pro-Social	HR	59%	187	5.401	0.120
	HH	74%	171	4.918	0.113
	HH-C	65%	142	4.824	0.130

Note: *Obs.* represents the number of observations (i.e. pairs round-subject).

We next test Hypothesis 4, which predicts that the task 1 duration of a pro-social Worker 1 is shorter in the HH and HH-C conditions than in the HR condition, while the task 1 duration of a selfish Worker 1 in the HH and HH-C conditions is not different from the HR condition. Table 6 pools the data from the HR, HH, and HH-C treatments and reports panel linear regressions of the task 1 duration on treatment dummies for the HH and HH-C conditions, separately for selfish and prosocial Workers 1. On the one hand, column (1) shows that the task 1 duration of selfish Workers 1 is not significantly different across treatments. On the other hand, column (2) shows that task 1's duration of pro-social Workers 1 is significantly shorter in the HH and HH-C treatments than in the HR treatment. These results support Hypothesis 4, providing evidence of the role of social preferences in the faster completion of task 1 when Worker 1 collaborates with another human.

Table (6) Treatment Effect on Task 1 Duration by Social Type: HH vs. HR vs. HH-C

	Task 1 Duration	
	Selfish	Pro-Social
	(1)	(2)
HH treatment	0.147	-0.498**
	(0.554)	(0.242)
HH-C treatment	-0.050	-0.571***
	(0.278)	(0.214)
Constant	5.415***	4.763***
	(0.400)	(0.581)
CRT	Yes	Yes
Round	Yes	Yes
Demographics	Yes	Yes
Observations	269	500
Subjects	29	56

Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. The table includes data from subjects playing the role of Worker 1 in the HH, HR and HH-C treatments. Significance reported: $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

Table (7) Task 2 Duration: HH and HH-C

	Task 2 Duration			
	HH	HH-C Treatment		
	(1)	(2)	(3)	(4)
Task 1 Duration	-0.173***	-0.185***	-0.180***	-0.181***
	(0.034)	(0.060)	(0.062)	(0.064)
Worker 2's Commitment	-	-	0.054	-
	-	-	(0.033)	-
Worker 2's Commitment to Finishing Fast	-	-	-	-0.146
	-	-	-	(0.103)
Constant	4.714***	4.284***	3.985***	4.319***
	(0.939)	(0.723)	(0.638)	(0.726)
CRT	Yes	Yes	Yes	Yes
Round	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Observations	235	219	205	205
Subjects	27	26	25	25

Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. Column (1) considers data from the HH treatment, while columns (2) to (4) consider data for the HH-C treatment. In column (3) we consider all data, while in columns (3) and (4) we exclude cases in which Worker 2 did not commit to complete task 2. Significance reported: $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

Finally, we note that while a main driver of Worker 1's decision to work faster when collaborating with another human is a concern for the collaborator's well-being, working faster is also beneficial for Worker 1, as it makes Worker 2 work faster as well. Table 7 presents regressions of the task 2 duration on the duration of task 1. Column (1) corresponds to the HH treatment, while

columns (2) to (4) correspond to the HH-C treatment (column (2) includes all the observations, while columns (3) and (4) include only observations where Worker 2 chose to commit and control for Worker 2's commitment as a continuous and binary variables, respectively). We observe that task 2 duration is significantly correlated with the task 1 duration in both the HH and HH-C treatments, suggesting that working faster when collaborating with a human Worker 2 is beneficial for Worker 1.

6. Conclusions

Motivated by the recent growth in applications whereby workers collaborate closely with robots/algorithms, we study the effects of process automation on workers' productivity in a setting in which workers must complete their tasks sequentially. We develop a theoretical model that captures the key aspects of a collaborative project: Workers receive a reward for completing the project and a bonus for early completion, and there is an alternative use of their time that provides them with a private payoff but does not contribute toward the completion of the project. We show that workers should always complete the project, and if the early completion bonus is high enough, they should do so as fast as possible. We conjecture that human workers deviate from the prescribed optimal policy and tend to delay the task. Process automation, in contrast, should lead to a higher project completion rate, shorter project completion time, and higher productivity of Worker 1.

To test these predictions and study the behavioral drivers of workers' decisions, we design an incentivized behavioral experiment that replicates the theoretical setting. The experimental results show that workers significantly depart from the optimal policy, as they take longer to complete their tasks or do not complete the project. While process automation leads to a higher project completion rate and shorter project duration, surprisingly, workers who collaborate with a robot take longer to complete their tasks, contradicting our initial hypothesis.

One possible explanation for this result is that, because workers know that the robot will complete the subsequent task with certainty, they do not need to provide the "wiggle room" that they would provide a human coworker who may be prone to delay the task. We further explore the effects of reduced uncertainty in regard to the completion of a subsequent task by using an additional treatment that reduces the uncertainty while keeping the collaboration between two human workers. We observe that reduced uncertainty has a more positive effect on a worker's productivity in the human-human setting than in the human-robot setting. Social preferences provide a good explanation for this result: Prosocial workers with other-regarding preferences work faster when collaborating with a human coworker than when they work alongside a robot.

Taken together, these results suggest that automation can be a double-edged sword, as it may increase the project completion rate and reduce the project completion time, but it also can decrease the productivity of workers who care about their coworkers' well-being.

6.1. Managerial Implications

Our results provide valuable insights for organizations that seek to automate parts of their processes. On the one hand, our results confirm that automation can be valuable. We observe that the completion time and completion rate improve if part of the process is automated. Hence, automation may be beneficial when these are the most relevant outcomes. Moreover, our results show that workers in charge of the early tasks tend to delay their tasks significantly longer than do those in charge of later ones. Therefore, if a fair division of work is the primary concern, it may be useful to automate one of the tasks to avoid these disparities. On the other hand, our results also show that automation can negatively affect workers' productivity and effort, especially among those who care about their coworkers' well-being. This observation indicates that the marginal benefit of the investment in process automation may be lower than the prediction in a theoretical benchmark. Although it is beyond the scope of our paper, we also note that the fact that workers take longer to complete their task when the subsequent task is automated is not necessarily an undesirable outcome. Firms may benefit in the long term from providing an improved, less stressful, work environment for their workers. In fact, citing a previous example, Chipotle's CTO reported that, by implementing robotic solutions, the firm seeks to create an easier, more fun, and more rewarding working environment, taking away from workers tasks that they do not enjoy doing and giving them more time to focus on tasks that they do enjoy (Rogers 2022).

We believe that more research is needed to further understand the behavioral effects of process automation. It would be valuable to study whether our main insights hold when early tasks are automated and to explore the impact when only a part of a worker's task is automated. Moreover, it would be useful to determine whether our experimental findings translate to more complex settings where multiple workers collaborate toward completing a project. On the one hand, the negative externalities of process automation may propagate over the entire process, and thus all workers may decrease their productivity if one task is automated. On the other hand, having multiple coworkers may preserve the positive impact of social preferences on workers' productivity and ameliorate the negative behavioral impact of process automation.

References

Andreoni J, Miller J (2002) Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica* 70(2):737–753.

Andreoni J, Vesterlund L (2001) Which is the fair sex? gender differences in altruism. *Quart. J. Econom.* 116(1):293–312.

Bai B, Dai H, Zhang DJ, Zhang F, Hu H (2022) The impacts of algorithmic work assignment on fairness perceptions and productivity: Evidence from field experiments. *Manufacturing & Service Operations Management* .

Baucells M, Grushka-Cockayne Y, Hwang W (2018) Managerial mental accounting and downstream project decisions. *Available at SSRN 3265724* .

Bearden JN, Murphy RO, Rapoport A (2008) Decision biases in revenue management: Some behavioral evidence. *Manufacturing Service Oper. Management* 10(4):625–636.

Beer R, Ahn HS, Leider S (2018) Can trustworthiness in a supply chain be signaled? *Management Sci.* 64(9):3974–3994.

Beer R, Qi A (2021) To communicate or not? Inter-firm communication in collaborative projects. *Available at SSRN 3290267* .

Beer R, Rios I, Saban D (2021) Increased transparency in procurement: The role of peer effects. *Management Sci.* 67(12):7511–7534.

Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. *Games Econom. Behavior* 10(1):122–142.

Bhaskaran SR, Krishnan V (2009) Effort, revenue, and cost sharing mechanisms for collaborative new product development. *Management Sci.* 55(7):1152–1169.

Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *Amer. Econom. Rev.* 90(1):166–193.

Cao X, Zhang D (2020) The impact of forced intervention on AI adoption. *Available at SSRN 3640862* .

Caro F, Colliard JE, Katok E, Ockenfels A, Stier-Moses N, Tucker C, Wu D (2022) Call for papers-*Management Science* special issue on the human-algorithm connection. *Management Sci.* 68(1):7–8.

Caro F, Saez de Tejada Cuenca A (2022) Believing in analytics: Managers' adherence to price recommendations from a DSS. *Manufacturing & Service Operations Management, forthcoming* .

Chen S, Klastorin T, Wagner MR (2021) Designing practical coordinating contracts in decentralized projects. *Naval Res. Logist.* 68(2):183–198.

Christin A (2017) Algorithms in practice: Comparing web journalism and criminal justice. *Big Data & Society* 4(2):2053951717718855.

Crama P, Sting FJ, Wu Y (2019) Encouraging help across projects. *Management Sci.* 65(3):1408–1429.

Cui T, Raju JS, Zhang ZJ (2007) Fairness and channel coordination. *Management Sci.* 53(8):1303–1314.

Davenport TH, Ronanki R (2018) Artificial intelligence for the real world. *Harvard Business Rev.* 96(1):108–116.

Davis AM, Huang R, Thomas DJ (2022) Retailer inventory sharing in two-tier supply chains: An experimental investigation. *Management Sci.* .

Dawande M, Janakiraman G, Qi A, Wu Q (2019) Optimal incentive contracts in project management. *Production Oper. Management* 28(6):1431–1445.

DHL (2022) Bring on the bots! How automation can break through bottlenecks. `https://www.dhl.com/global-en/delivered/digitalization/logistics-automation.html`, retrieved November 22, 2022.

Diakopoulos N (2020) Accountability, transparency, and algorithms. *The Oxford Handbook of Ethics of AI* 17(4):197.

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144(1):114.

Dietvorst BJ, Simmons JP, Massey C (2018) Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science* 64(3):1155–1170.

Dohmen T, Falk A (2011) Performance pay and multidimensional sorting: Productivity, preferences, and gender. *Amer. Econom. Rev.* 101(2):556–90.

Duch ML, Grossmann MR, Lauer T (2020) z-tree unleashed: A novel client-integrating architecture for conducting z-tree experiments over the internet. *Journal of Behavioral and Experimental Finance* 28:100400.

Duffy J, Li Y (2019) Lifecycle consumption under different income profiles: Evidence and theory. *J. Econom. Dynamic Control* 104:74–94.

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.

Frederick S (2005) Cognitive reflection and decision making. *J. Econom. Perspectives* 19(4):25–42.

Ge R, Zheng Z, Tian X, Liao L (2021) Human–robot interaction: when investors adjust the usage of robo-advisors in peer-to-peer lending. *Information Systems Research* 32(3):774–785.

Grushka-Cockayne Y, Erat S, Wooten J, Donohue K, Katok E, Leider S (2018) New product development and project management decisions. *The Handbook of Behavioral Operations*, 367–392 (Wiley).

Ho TH, Su X, Wu Y (2014) Distributional and peer-induced fairness in supply chain contract design. *Production Oper. Management* 23(2):161–175.

Hou C, Lu M, Deng T, Shen ZJM (2021) Coordinating project outsourcing through bilateral contract negotiations. *Manufacturing Service Oper. Management* 23(6):1543–1561.

Ibrahim R, Kim SH, Tong J (2021) Eliciting human judgment for prediction algorithms. *Management Sci.* 67(4):2314–2325.

Kagan E, Leider S, Sahin O (2021) Dynamic decision-making in operations management. *Available at SSRN* .

Kalkancı B, Chen KY, Erhun F (2014) Complexity as a contract design factor: A human-to-human experimental study. *Production Oper. Management* 23(2):269–284.

Katok E, Pavlov V (2013) Fairness in supply chain contracts: A laboratory study. *J. Oper. Management* 31(3):129–137.

Kim SH, Tong J, Peden C (2020) Admission control biases in hospital unit capacity management: How occupancy information hurdles and decision noise impact utilization. *Management Sci.* 66(11):5151–5170.

Knight W (2014) How human-robot teamwork will upend manufacturing. `https://www.technologyreview.com/2014/09/16/171369/how-human-robot-teamwork-will-upend-manufacturing/`, retrieved November 22, 2022.

Kwon HD, Lippman SA, McCardle KF, Tang CS (2010) Project management contracts with delayed payments. *Manufacturing Service Oper. Management* 12(4):692–707.

Lebovitz S, Lifshitz-Assaf H, Levina N (2022) To engage or not to engage with ai for critical judgments: How professionals deal with opacity when using ai for medical diagnosis. *Organization Science* 33(1):126–148.

Leider S, Lovejoy WS (2015) Bargaining in supply chains (long version). Working paper.

Leider S, Şahin Ö (2014) Contracts, biases, and consumption of access services. *Management Sci.* 60(9):2198–2222.

Li J, Leider S, Beil D, Duenyas I (2021) Running online experiments using web-conferencing software. *Journal of the Economic Science Association* 7(2):167–183.

Lin W, Kim SH, Tong J (2022) What drives algorithm use? An empirical analysis of algorithm use in type 1 diabetes self-management. *Available at SSRN 3891832* .

Long X, Nasiry J, Wu Y (2016) A behavioral study on abandonment decisions in multistage projects. *Management Sci.* 66(5):1999–2016.

Moffitt KC, Rozario AM, Vasarhelyi MA (2018) Robotic process automation for auditing. *J. Emerging Technologies in Accounting* 15(1):1–10.

Olsen TL, Tomlin B (2020) Industry 4.0: Opportunities and challenges for operations management. *Manufacturing Service Oper. Management* 22(1):113–122.

Palley AB, Kremer M (2014) Sequential search and learning from rank feedback: Theory and experimental evidence. *Management Sci.* 60(10):2525–2542.

Rahmani M, Roels G, Karmarkar US (2017) Collaborative work dynamics in projects with co-production. *Production Oper. Management* 26(4):686–703.

Rogers K (2022) Chipotle is testing a robot that makes tortilla chips — and its name is 'Chippy'. `https://www.cnbc.com/2022/03/16/chipotle-tests-robot-kitchen-assistant-named-chippy-to-make-tortilla-chips.html`, retrieved November 22, 2022.

Russ H (2022) Starbucks projects profit growth from tech, stores, workers spending. `https://www.reuters.com/business/retail-consumer/starbucks-exploring-tech-ease-high-us-demand-worker-stress-2022-09-13/#:~:text=0)%20projects%20profits%20to%20grow,coffee%20chain%20said%20on%20Tuesday.`, retrieved November 22, 2022.

Shokoohyar S, Qi A, Katok E (2019) Improving coordination in assembly projects. *Available at SSRN 2889868* .

Siemsen E, Aloysius J (2020) Supply chain analytics and the evolving work of supply chain managers. *Association for Supply Chain Management White Paper* .

Snyder C, Keppler S, Leider S (2022) Algorithm reliance under pressure: The effect of customer load on service workers. *Available at SSRN 4066823* .

Song JM, Zhao Y, Xu X (2021) Incentives and gaming in collaborative projects under risk-sharing partnerships. *Manufacturing Service Oper. Management* 23(2):453–470.

Sun J, Zhang DJ, Hu H, Van Mieghem JA (2022) Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations. *Management Science* 68(2):846–865.

Wu Y, Ramachandran K, Krishnan V (2014) Managing cost salience and procrastination in projects: Compensation and team composition. *Production Oper. Management* 23(8):1299–1311.

Wuttke DA, Donohue K, Siemsen E (2018) Initiating supplier new product development projects: A behavioral investigation. *Production Oper. Management* 27(1):80–99.

Zhang J (2016) Deadlines in product development. *Management Sci.* 62(11):3310–3326.

Zhao S, López Vargas K, Friedman D, Gutierrez M (2020) Ucsc leeps lab protocol for online economics experiments. *Available at SSRN 3594027* .

Appendix A: Proofs

Proof of Proposition 1. We first note that when W_2's task is not active, i.e., $\mathbb{1}_{\{r_1=0\}}=0$, W_2's only feasible decision is to take the alternative option, i.e., $x_1^*=0$. In the proof below, we discuss the scenario when W_2's task is active. That is, W_1 has completed her task, i.e., $r_1=0$. In this case, W_1's only feasible decision is to take the alternative option, i.e., $x_2^*=0$, and the continuation value of W_2 is as follows:

$$\begin{aligned} V_{2,t}((0,r_2),\mathbf{a}) &= \max_{x_2\in\{0,1\}}\left\{(1-x_2)\cdot a_2+\mathbb{1}_{\{r_2=0\}}\cdot\delta_2+\mathbb{E}_{\mathbf{a}_{t+1}}\left[V_{2,t+1}\left((0,r_2-x_2),\mathbf{a}_{t+1}\right)\right]\right\} \\ V_{2,T+1}((0,r_2),\mathbf{a}) &= \pi_2\cdot\mathbb{1}_{\{r_2=0\}}. \end{aligned} \tag{6}$$

Before we proceed, we make the following observations. First, it is easy to see that if W_2 has completed her task, i.e., $r_2=0$, then W_2 does not have incentives to work on the project, i.e., $x_2^*=0$. It follows that if $r_2=0$ in period $t\le T$, we have that

$$V_{2,t}((0,0),\mathbf{a})=\pi_2+\delta_2\cdot(T-t+1)+a_2+\mathbb{E}[a]\cdot(T-t). \tag{7}$$

Second, it is also easy to see that if W_2 does not have enough time to complete her task, i.e., $r_2>T-t+1$, then W_2 is better off taking the alternative options instead of working on the project in the remaining periods. It follows that if $r_2>T-t+1$, we have that

$$V_{2,t}((0,r_2),\mathbf{a})=a_2+\mathbb{E}[a]\cdot(T-t). \tag{8}$$

We next establish the optimal policy of W_2 in period t for $0<r_2\le T-t+1$ by induction.

- Base: Consider $t=T$. If $r_2=1$, we have $x_2^*=1$, since if W_2 chooses the alternative option in this period (i.e., $x_2=0$), her payoff-to-go is a_2, and if W_2 chooses to work on the project (i.e., $x_2=1$), her payoff-to-go is π_2. By Assumption 1, we have that $\pi_2\ge h+\mathbb{E}[a](\tau_2-1)\ge h\ge a_2$.
- Induction: suppose that the optimal decision of W_2 is given by (2) if $\delta_2\ge h-\mathbb{E}[a]$, or by (3) if $\delta_2<h-\mathbb{E}[a]$ for periods $\{t+1,\ldots,T\}$. We now show that it is also true in period t for $0<r_2\le T-t+1$.
  - When $r_2=T-t+1$, if W_2 chooses to take the alternative option in the period, she will not have enough time to complete the project and her payoff is $a_2+\mathbb{E}[a]\cdot(T-t)$ by (8). If W_2 chooses to work on the project in the period, then she will keep working on the project until the end of the time horizon and her payoff is π_2. By Assumption 1, we have that $\pi_2\ge h+\mathbb{E}[a]\cdot(\tau_2-1)\ge a_2+\mathbb{E}[a]\cdot(T-t)$. That is, W_2 should work on the project in period t.

— When $0 < r_2 < T - t + 1$, if W_2 chooses to take the alternative option, then her payoff is

$$a_2 + \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right]. \tag{9}$$

If W_2 chooses to work on the project, then her payoff is

$$\mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right]. \tag{10}$$

First, when $a_2 = h$ and $\delta_2 < h - \mathbb{E}[a]$, we have that:

$$\begin{aligned} h + \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right] &\geq h - \max\{h, \delta_2 + \mathbb{E}[a]\} + \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right] \\ &\geq \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right]. \end{aligned}$$

The first inequality follows from the fact that when there are r_2 remaining periods of task at the beginning of period $t+1$, if W_2 instead uses the optimal strategy with $r_2 - 1$ remaining periods of task at the beginning of period $t+1$, with the modification that W_2 should work on the project in the first period when W_2 would have taken the alternative option, then it is easy to see that $\mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right] \geq -\max\{h, \delta_2 + \mathbb{E}[a]\} + \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right]$. The second inequality follows from $\delta_2 < h - \mathbb{E}[a]$. Therefore, W_2 is better off taking the alternative option in this case.

Second, when $a_2 = l$ or $\delta_2 \geq h - \mathbb{E}[a]$, we have that:

$$\begin{aligned} \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right] &\geq \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right] + \delta_2 + \mathbb{E}[a] \\ &\geq \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right] + a_2. \end{aligned}$$

The first inequality follows from the fact that when there are $r_2 - 1$ remaining periods of task at the beginning of period $t+1$, if W_2 instead uses the optimal strategy with r_2 remaining periods of task at the beginning of period $t+1$, with the modification that W_2 should take the alternative option in the last period when W_2 would have worked on the project, then it is easy to see that $\mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2 - 1), \mathbf{a}_{t+1} \right) \right] \geq \mathbb{E}_{\mathbf{a}_{t+1}} \left[V_{2,t+1} \left((0, r_2), \mathbf{a}_{t+1} \right) \right] + \delta_2 + \mathbb{E}[a]$. The second inequality follows from either $\mathbb{E}[a] \geq l$ or $\delta_2 \geq h - \mathbb{E}[a]$. Therefore, W_2 is better off working on the project in this case. ■

Proof of Proposition 2. The proof is similar to the proof of Proposition 1 with the effective deadline of W_1 changed from T (for W_2) to $T - \tau_2$ and is therefore omitted for brevity. ■

Appendix B: Appendix to the Experimental Design Section

B.1. Main Game Instructions - HH Treatment

Welcome and thank you all for coming. Today's session is a study of individual decision making in the context of collaboration between workers. The study takes approximately 75 minutes to complete. Please pay close attention to the instructions. To ensure that you understand the instructions, we will ask you several questions as we go along. You will only be allowed to proceed to the task if you can answer those questions. If you complete this experiment you will receive a participation payment of $5. In addition, you will receive a bonus payment. Typically bonus payments range between $5 and $20 and depend on your and other participants' decisions and on chance. To continue, please review the consent form in the next screen.

You will first play the Product Launch Game we will now describe to you. In this game you will make decisions that will earn you points. Your points will be converted to dollars at a rate of 50 points = 1 dollar, and you will be paid in dollars via an Amazon gift card. You will be asked to complete a separate form to claim your payment at the end of the experiment You will then play four short additional games we will later describe to you. In these games, you will also earn points which will be converted to dollars for payment. Your payoff for today's session will be the $5 show up fee plus your dollar earnings from all the games. If you have a question at any time, please use the chat to enter your question and the experimenters will answer it. We ask that you do not chat with one another for the duration of the experiment. Also, please put aside any material that is not related to the experiment.

The Product Launch Game. You will first play ten rounds of the Product Launch Game. At the beginning of the game, each of you will be randomly assigned a role, "Worker 1" or "Worker 2", and you will keep the same role for the ten rounds of the game. You will be matched with a different participant for each round of the game and you will not know which participant you are matched with. The participant you are matched with will have the opposite role as you have. Worker 1 and Worker 2 are working together in a project to develop a new product. The project consists of two sequential tasks, 1 and 2, which are assigned to Worker 1 and 2, respectively. Worker 1 can start working on task 1 in period 1, and this task must be completed before Worker 2 can start working on task 2. The project is completed when task 2 is completed. At that point, the product is launched. The deadline to complete the project is period 12.

Worker 1 needs to spend 3 periods of work to complete task 1, and Worker 2 needs to spend 3 periods of work to complete task 2. Recall that Worker 2 cannot start working on the project until Worker 1 has

completed task 1. If the project is completed before the deadline, each worker earns 100 points from the project completion. In addition, for each period the project is finished early, each worker earns 30 points. For example, if the project is completed in period 8 (that is, 4 periods before the period 12 deadline), each worker earns an additional early termination bonus of 4 periods $\times$ 30 points/period $=$120 points.

In each period, if a worker's task is active (that is, if the worker can already work on the task and that task has not been completed yet) the worker can choose to work on the task or to spend the period on an alternative option that gives the worker a private benefit but does not contribute to the value of the project. When a worker's task is not active (either because they cannot start working on it yet, or because the task is already completed) the worker earns the value of the alternative option. When a worker works on the project, she does not earn the value of the alternative option. The value of the alternative option is random in each period, and it is different for the two workers. It takes a value of 40 points or 10 points with equal chance.

How to earn money. Workers' payoffs are computed as:

$$\text{Worker's Payoff} = [\text{project completion payment}] + [\text{early completion bonus}] + [\text{sum of alternative options}],$$

where

1. Project completion payment: If the project is completed before the deadline = 100 points. Otherwise, = 0 points.
2. Early completion bonus: If the project is completed X periods early (before the deadline) $= X$ periods $\times$ 30 points/period. Otherwise, = 0 points.
3. Sum of alternative options: The sum of the value of the alternative option earned in each period where the worker did not work on the project.

At the end of each round everyone will be randomly re-matched for the following round. You will not know which subjects you are matched with. At the end of the experiment, one of the ten rounds of the Product Launch Game will be randomly selected for payment. Your points in that round will be converted to dollars at a rate of 50 points = 1 dollar. In addition, at the end of the experiment you will participate in four short additional games. You will receive the instructions for these additional games once the Product Launch Game is over. You will be paid the $5 show up fee plus your earnings from the Product Launch Game and the other additional games.

B.2. CRT Game

The questions included in the CRT game are the following:

1. If it takes 5 machines to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets?
2. In a lake, there is a patch of lily pads. every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days it takes for the patch to cover half of the lake?
3. A sandwich and a drink cost $5.50 in total. The sandwich costs $5.00 more than the drink. How many cents does the drink cost?
4. Each month, Rolando earns a commision of 10% of his sales for the month, plus a salary of $2500. If Rolando earns $3000 in a certain month, what were his total sales?

B.3. Interface Screenshots

Figure (2) Main Game Period 4 Decision Screen: HH

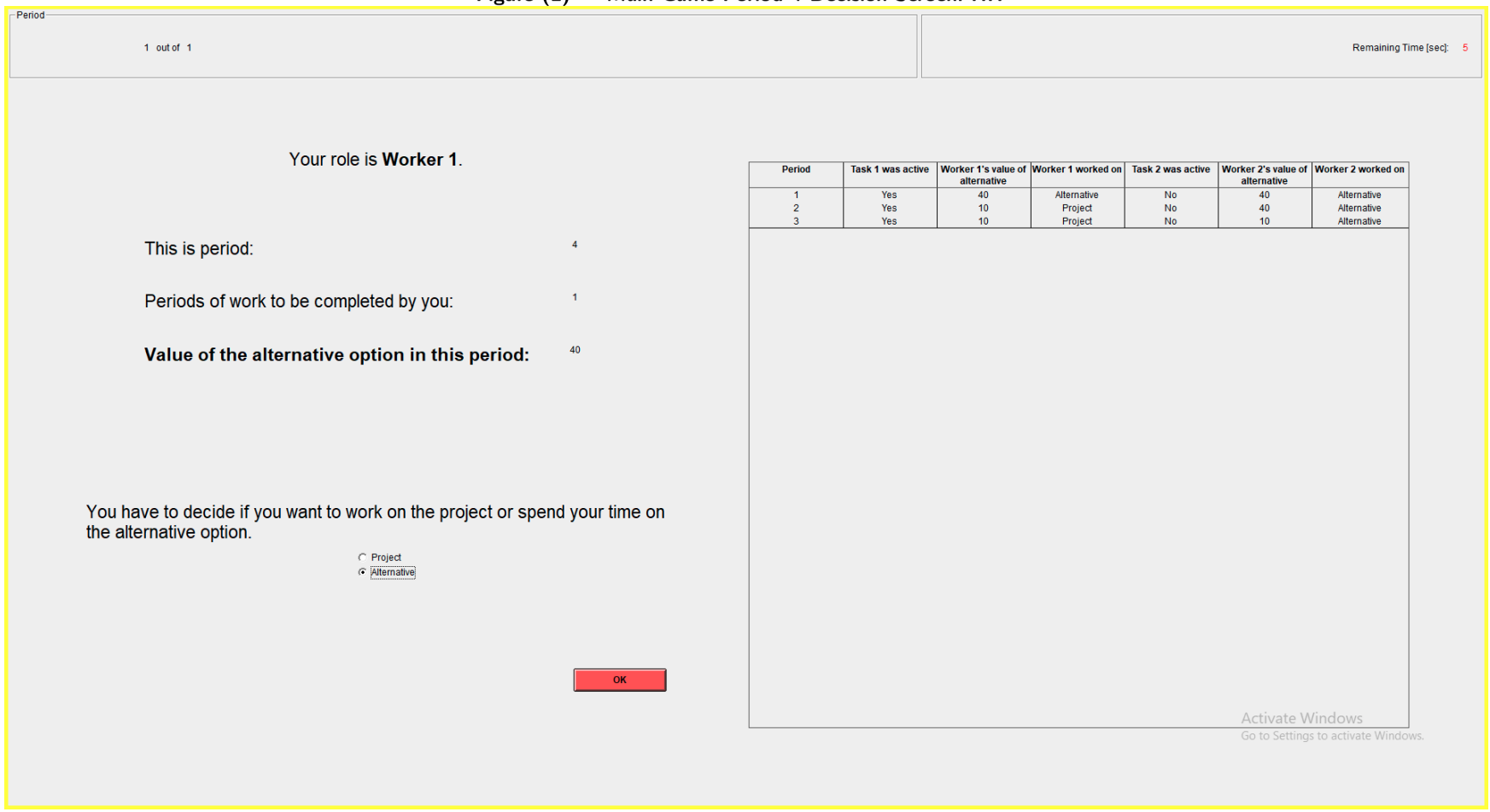
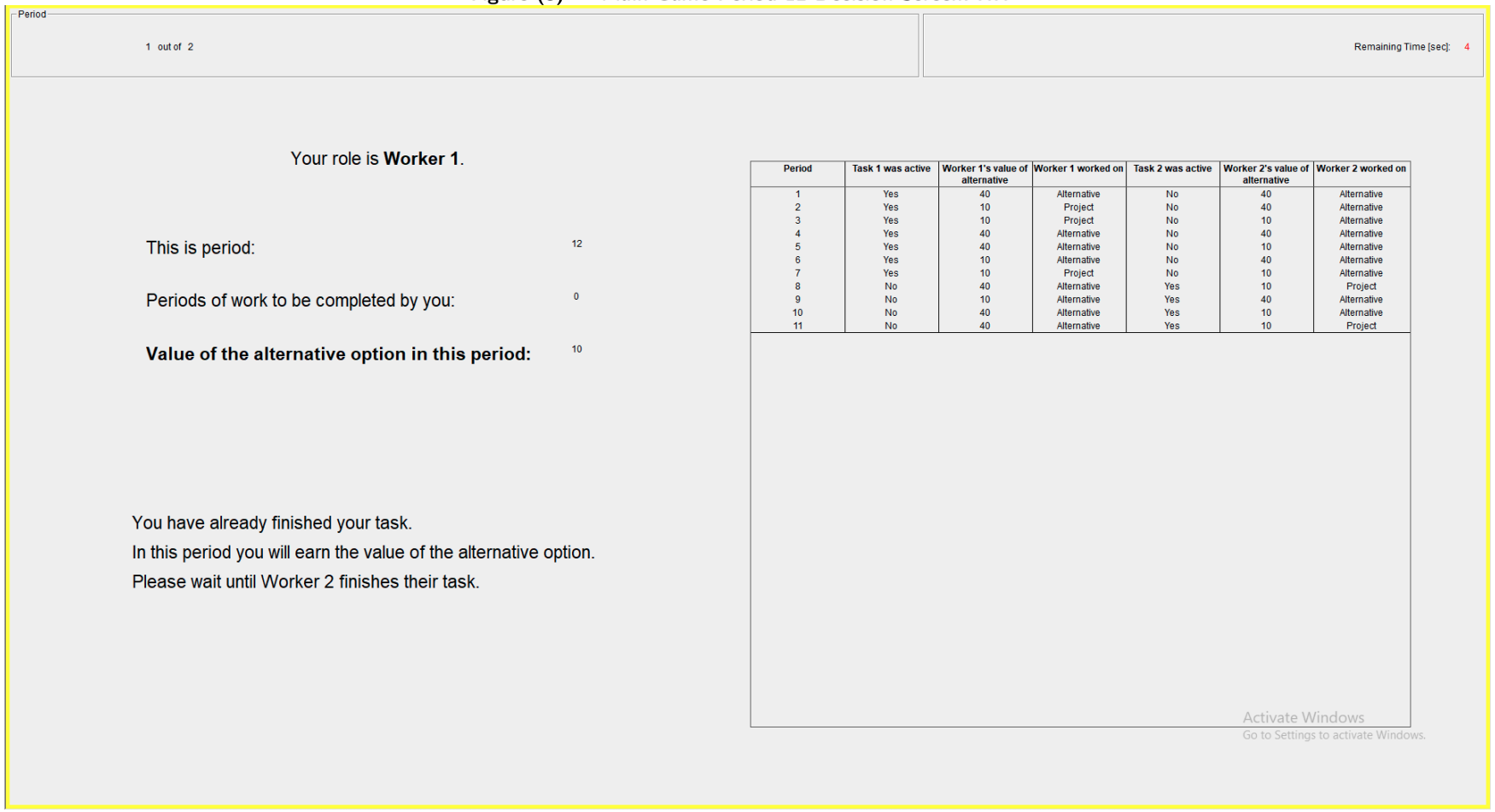


Figure (3) Main Game Period 12 Decision Screen: HH



Period	Task 1 was active	Worker 1's value of alternative	Worker 1 worked on	Task 2 was active	Worker 2's value of alternative	Worker 2 worked on
1	Yes	40	Alternative	No	40	Alternative
2	Yes	10	Project	No	40	Alternative
3	Yes	10	Project	No	10	Alternative
4	Yes	40	Alternative	No	40	Alternative
5	Yes	40	Alternative	No	10	Alternative
6	Yes	10	Alternative	No	40	Alternative
7	Yes	10	Project	No	10	Alternative
8	No	40	Alternative	Yes	10	Project
9	No	10	Alternative	Yes	40	Alternative
10	No	40	Alternative	Yes	10	Alternative
11	No	40	Alternative	Yes	10	Project

Appendix C: Appendix to the Experimental Results Section

C.1. Self-Reported Strategies.

In this appendix, we analyze subjects' comments in the exit survey to get further insights. While self-reported strategies may be subject to biases and thus should be considered with caution, we observed that the comments provide additional validation to some of our assumptions and results.

First, subjects' comments confirm that the values of the parameters chosen in the experiment create a setting where they may be tempted to deviate from the optimal policy and delay the project when the realized value of the outside option is high (e.g., one subject commented: *"...I was still experimenting to find the right strategy to complete the project as well as earn maximum reward. However there were few instances were the high reward tempted me to choose the alternative which actually felt like a trap to earn less..."*).

Second, a key distinction between the HH/HH-C and the HR treatments is the uncertainty about task 2's duration and the project completion in the HH/HH-C cases. We conjectured that in the conditions where both tasks were completed by human workers, Worker 1 experienced *uncertainty* about whether Worker 2 would complete her task (and, therefore, about whether the project would be completed). In line with the conjecture that uncertainty is a key driver of behavior, a subject commented: *"I was trying to finish my task by period 5 because I realized that, when I took more time than that, the Worker 2 tended to not complete their task."*

Finally, several of the subjects' comments also reinforce our findings regarding social preferences as a behavioral driver behind Worker 1's faster task completion when collaborating with another human vs. the automated robot. We propose that two distinct forms of social preferences may be at play: first, Worker 1 may perceive that completing her task fast is *fair* to Worker 2, as it gives Worker 2 more time to complete his own task. Second, Worker 1 may be prone to complete her task fast out of concerns for *surplus maximization*, as both workers benefit from the early completion bonus. The subjects' comments reflected that both forms of social preferences are present. For example, a subject indicated that *"I completed the project as soon as possible. I did this because it would give my partner the most time to complete the project"*, suggesting preferences for fairness. Consistent with preferences for surplus maximization, a subject commented: *"I was trying to get the same points for the Worker 2 as well"*, while another subject commented: *"My strategy was to balance between the alternative and getting the (early completion) bonus so that both team members get a decent amount of money"*.

C.2. Social Preferences.

In the Altruism Game, subjects face 5 situations in which they have to allocate 50 tokens between themselves and a partner. Situations differ in the value that each token represents for the sender and the receiver. Specifically, in situations 1 to 5 the values for the sender are $u = \{0.15, 0.1, 0.05, 0.05, 0.05\}$, while the values for the receiver are $v = \{0.05, 0.05, 0.05, 0.1, 0.15\}$.

Based on the decisions of each subject in the five scenarios, we compute several metrics:

- Own Allocation: We compute the minimum and maximum tokens allocated to themselves across situations.
- Fairness: for each situation, we compute the absolute value of the difference in earnings in that situation and then compute the minimum and maximum across situations.
- Surplus: for each situation, we compute the surplus (the total earning of the two parties) generated by the allocation and then compute the maximum surplus.[11] In addition, we compute the difference between the surplus generated in the first and last situations (i.e., we compare the most favorable situation for the sender and the receiver in terms of value of each token).

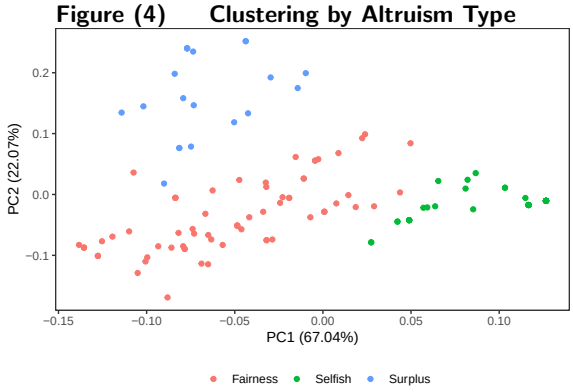
Based on the computed metrics, we cluster the subjects into three groups:

1. "Selfish" if they show a tendency to keep all the tokens,
2. "Fairness maximizers", if they show a tendency to equalize payoffs,
3. "Surplus maximizers" if they show a tendency to maximize the sum of their own and the other subject's payoffs.

To validate our clustering results, we first conduct a Principal Component Analysis and show the three types based on the first two principal components (accounted for 67.04% and 22.07% percentages of explained variances respectively) in Figure 4. It is clear that the Selfish (green dots), Fairness maximizers (red dots), and Surplus maximizers (blue dots) types are separated from each other. Note that the observations could be overlapping since they have the same values for the principal components, and therefore, the number of dots does not necessarily reflect the number of observations.

[11] Notice that we do not compute the minimum surplus because it is always the same (obtained from the third situation).

Figure (4) Clustering by Altruism Type



Second, we report summary statistics for the main variables of interest, separated by group, in Table 8. We make the following observations. First, we observe that subjects in the Selfish group allocate most tokens to themselves, which leads to the most unfair allocations. Second, we find that subjects in the Fairness group avoid extreme allocations of tokens. Finally, we observe that subjects in the group of Surplus maximizers achieve the lowest value of the difference between the surplus generated in the first and last situations and a large values of the maximum surplus, which is consistent with the fact that they try to maximize the overall surplus in all situations. Overall, we conclude that the results in Table 8 correctly characterize the three aforementioned groups.

Table (8) Summary Statistics by Altruism Type

		Own Allocation		Fairness		Surplus	
Group	N	Min.	Max.	Min.	Max.	Max	Diff. Ext.
Selfish	49	46.245	48.388	2.032	6.969	7.235	4.555
Surplus	20	5.400	46.400	0.692	6.400	6.950	0.380
Fairness	69	22.812	39.565	0.274	3.730	5.571	1.652

Note: Each observation corresponds to one subject.

C.3. Additional Results.

Table (9) General Results: HH vs. HH-C

	Project Completion	Project Duration	Task 1 Duration	Task 2 Duration
HH-C Treatment	-0.022	-0.359	-0.099	-0.252
	(0.048)	(0.314)	(0.180)	(0.245)
Constant	0.896***	8.311***	3.909***	3.893***
	(0.056)	(0.372)	(0.394)	(0.533)
CRT	Yes	Yes	Yes	Yes
Round	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Observations	530	454	454	454
Subjects	53	53	53	53

Note: Panel linear regressions with subject random effects. Standard errors clustered at the session level reported in parentheses. The table pools data from subjects playing the role of Worker 1 in the HH and HH-C treatments. Significance reported: $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

Figure (5) Commitment of Worker 2 (cond. on completed project): HH-C