FROM ME TO WE: BEATING PROCRASTINATION IN TEAMS

ANUJIT CHAKRABORTY* AND GUIDON FENIG**

ABSTRACT. Can team incentives increase worker's productivity and decrease procrastination in intertemporal tasks? We recruited just under 600 online workers to engage in tedious tasks over three days. They were randomly assigned to either individualistic (Solo) incentives or to one of two team-based incentives (Cooperative and Competitive). Contrary to theoretical predictions, workers under Cooperative incentives surpassed the performance of those working under either Solo or Competitive incentives. Productivity on Day 1, which in theory should inversely relate to procrastination, was also significantly higher in both team treatments. Our structural analysis confirms that teams increase productivity by enhancing intrinsic motivation *and* by reducing the tendency to delay work. Finally, teams increase productivity further under Competitive incentives, when workers can observe and react to the efforts of their team members.

JEL classifications: C9, C72, C92, D9.
Keywords: intertemporal tasks, time preferences, present bias, intrinsic motivation, real effort tasks, team-based incentives

A vast literature in economics and psychology finds that individuals systematically procrastinate: we persistently postpone unpleasant but important activities, which ends up reducing our overall productivity. Although most of this empirical evidence on procrastination (see Frederick et al. 2002, Cohen et al. 2020 for a review) is exclusively based on data from individualistic tasks, one expects a fundamental tendency like procrastination to also extend, perhaps equally, to tasks performed as a collective. When couples within a household plan to buy a home or jointly save for retirement, when co-workers compete among themselves to get a raise, or when players in a team sport aim to secure a league victory, individuals function as members of a broader collective or team, collectively pursuing interconnected objectives. Are individuals

* Department of Economics, University of California, Davis. (e-mail chakraborty@ucdavis.edu).
** Department of Economics, University of Ottawa, Ontario. (e-mail gfenig@uottawa.ca).
Financial support from SSHRC is gratefully acknowledged. Jeremy Epp provided valuable research assistance. We thank David Freeman, David Gill, Yoram Halevy, Radovan Vadovič, and Lanny Zrill for their valuable feedback. Our experiment was approved under University of Ottawa REB S-08-20-6072.

equally procrastinating and (un)productive across individual and collective environments, or could interconnected team-based goals help individuals be more motivated and procrastinate less?

In this paper, we measure how team-based incentives impact the early and total performance of individuals engaged in a multi-day real-effort task. We recruit under 600 online workers to participate in a real-effort study involving a straightforward typing task conducted over three consecutive workdays. These workers are randomly assigned to treatment conditions that offered personal or team-based incentives.

We study two different team-based incentives that utilize two fundamental human traits: cooperation and competition. In our *Cooperative* treatment, workers in a team of two equally share the total revenue generated collectively by them, based on a piece rate of p. Such incentive structures, based on sharing, are prevalent in numerous workplace settings. For instance, at Great-Plains Software (later acquired by Microsoft in 2001), team members shared compensation equally.[1] Similarly in RR Donnelly & Sons, a large commercial printing company, workers were organized into teams and rewarded collectively for achieving waste reduction goals. The cash bonus for meeting the target was evenly distributed among the team members.[2] In sharp contrast, in our *Competitive* treatment, both workers in a team of two get a minimum guaranteed piece rate of $2p/3$, but the worker who completes more tasks among the pair earns an additional bonus payment that depends on their total productivity. This payment scheme emulates the competitive compensation packages offered by companies that incentivize high-performers with company stocks or bonuses proportionate to company profits. To compare with these two team-based incentives, we implement an individualistic (*Solo*) incentive treatment, where workers are paid based on their own effort by the piece rate p.

In all treatment conditions, participants were allowed to work for a maximum of 45 minutes over three consecutive days. Participants had full flexibility in determining their daily work amount, allowing them to postpone the majority of tasks until the deadline. Essentially, instead of facing an explicit choice between "effort now" and "effort later" commonly found in economics experiments (for example, Coller and Williams 1999 utilize price lists and Augenblick et al. 2015 utilize convex time budgets), our study's participants encountered an implicit trade-off between present and

[1] See Parker et al. 2000 for details. Barnes et al. 2011 describe Lotus Software employing similar team incentives

[2] See Baldwin et al. 2008) for details.

future effort. This trade-off was determined by the overall time constraint (45 minutes) and the completion deadline (on Day 3). For any fixed amount of total work (or total compensation) one aims to complete before the deadline, working less on earlier dates means having to work more on subsequent dates. Thus, the intertemporal tradeoff is easy to understand as it mirrors real-life scenarios of working towards a deadline (for e.g., Ariely and Wertenbroch 2002).

As we later explain in Figure 2 in Section I, the payment schemes are designed so that all three treatments *always provide equal compensation to the average worker*; making them comparable in terms of overall earnings. Still the incentives are biased against the team treatments, especially the Cooperative treatment where the effective piecerate was only half of the Solo treatment. Theoretically, Cooperative incentives are predicted to perform the worst, for any arbitrary beliefs about teammates, even after accounting for reasonable levels of intrinsic motivation.

Contrary to the theoretical predictions, Cooperative incentives consistently outperform Solo and Competitive incentives across both measures of worker productivity: the probability of completing all 45 minutes of work and the total tasks completed. Furthermore, the difference between Cooperative and Solo incentives is consistently statistically significant. For example, compared to Solo, workers in the Cooperative treatment exhibit a 18.4% higher likelihood of working for the entire 45 minutes and they complete 65 more typing tasks.

Our second theoretical result connects Day 1's effort to present bias or procrastination: the more present-biased a worker is, the less they would work on Day 1, choosing instead to postpone work to the later days. We find that workers in both team treatments work more on Day 1 than those in the Solo treatment. This is particularly evident in the Competitive treatment, where the increase in productivity over the three days is largely due to the additional work done on Day 1. In the Cooperative treatment, the rise in Day 1 productivity accounts for more than a third of the total productivity increase compared to the Solo treatment. Thus, workers are less likely to postpone work under team incentives.

Based on the patterns in the data, two non-exclusive mechanisms can explain why team incentives yield better outcomes compared to individual incentives. Firstly, team incentives have the potential to enhance intrinsic motivation among team members (Ryan and Deci, 2000), thus making them work harder for the same piece rate. The second *novel* mechanism involves team incentives reducing present bias, which in turn increases productivity. To assess and analyze the two mechanisms, as well as their combined effects, we conduct a structural estimation exercise. We find statistically significant evidence *supporting both mechanisms*. In our model comparison tests, the conjunction hypothesis of "team incentives decrease present-biased behavior and motivate more than the implied piece rate" outperforms each of the individual explanations.

In the baseline team treatments, workers could only infer their partner's effort from their final payment, received after the study had concluded. In the next two treatments, we also investigate two additional team-based mechanisms that might influence intertemporal allocation of effort *if* workers were also informed about their fellow team member's Day 1 and Day 2 effort on Days 2 and 3, respectively. Firstly, we anticipated that workers would exert more effort on Days 2 or 3 if they observed their partners working harder on preceding days, driven by reciprocity in one treatment and competitiveness in the other, respectively. We call this the *Reaction* effect. Secondly, closely tied to the Reaction channel, we expected team members to anticipate the Reaction effect and modify their actions to influence the Reaction of other team members in a manner that favors their own outcomes. For example, anticipating the Reaction, workers in the Cooperative treatment could work harder on Day 1 to signal their type, countering their procrastination: we call this the *Signaling* effect. To study the marginal effect of these channels, we introduce two *observability (obs)* treatments: Cooperative-obs and Competitive-obs, *where workers could observe their partner's past day's effort*. Essentially, we *decompose* the informational stimulus inherent in teams by differentiating between our baseline and the *obs* team treatments. The baseline captures the effect of ex-post information, while the *obs* treatments pinpoint the impact of real-time feedback, which can not only influence immediate reactions but also induce anticipatory signaling effects.

We find significant evidence of the Reaction effect in both Cooperative-obs and Competitive-obs treatments. However, this effect is much stronger under Competitive-obs incentives. For example, observability increases total productivity by 41 tasks on average under competitive incentives, which results in Competitive-obs significantly outperforming Solo incentives and marginally outperforming Cooperation-obs on aggregate productivity.

Surprisingly, despite the presence of the Reaction effect, we fail to find any evidence of its dual: the Signaling effect. Effort in the first period remains mostly unchanged irrespective of whether observability is present, and this is regardless of the incentive type. For example, workers in the Cooperative treatment do not work more on Day

1 to influence their partner's future effort. Therefore, contrary to our expectations, strategic thinking does not impact the Day 1 choices. To unravel this paradox, we introduce two additional diagnostic treatments (see Section C in the Appendix) that reveals that most workers anticipate the Reaction effect, yet they lack the strategic sophistication required to leverage their opponent's responses to their own advantage.

To the best of our knowledge, this is the first paper to show that team incentives can also influence the intertemporal allocation of effort, a phenomenon generally attributed to *immutable* time preferences. Our results help explain why team incentives are being identified as powerful motivators in driving positive health outcomes, especially in activities prone to procrastination, for example losing weight or exercising. For example, Patel et al. 2016 study the percentage of subjects who meet a daily 7000-steps goal under individual incentives ($50 if subjects meet goal) versus under team incentive ($50 only if all four team members met the goal), and find that the latter was significantly higher. Similarly, Pearson et al. 2020 found that adding modest team-based incentives to an app-based rewards program significantly increased mean daily step count. Finally, Babcock et al. 2015 compares individual incentives (where a participant earns a $25 bonus for n visits to the gym or to the library, respectively in two separate experiments)with team incentives (a $25 bonus awarded only if both the subject and their partner attend n times). They find that team incentives significantly outperform individual ones.

Instead of the intertemporal patterns, most of the literature on teams so far has instead focused on how teams increase *aggregate productivity,* or how team incentives are affected by *social-preferences* and *social connections*. For example, in a series of influential papers on workplace incentives among teams of fruit pickers, Bandiera et al. [2005, 2010, 2013] investigate the relationship between social connections and social incentives, and how those affect total productivity. Similar connections have also been studied in broader contexts: researchers have asked subjects to work for charities (Imas 2014), induced a gift-exchange between employer and employee (Gneezy and List 2006), allowed social comparison within agents (Goldstein et al. 2008), or created a weakest-link payoff structure that demands coordination (Babcock et al. 2015, Bortolotti et al. 2016, List and Shah 2022) and found that such interventions increased producitivity. Even though some of these studies, especially the ones run in a field setting, use a multi-day design to compare aggregate effort across treatments, they do not focus on how that aggregate productivity was distributed intertemporally, as we do. Overall, our paper lies at the intersection of three experimental literatures.

The first strand measures the intertemporalallocation of *real effort* in individualistic decision tasks to parametrically estimate time preference parameters at the individual level (see Cohen et al. 2020 for a review). Secondly, our work is also related to the literature that compares aggregate productivity under different social compensation schemes (e.g., Bandiera et al. 2005, Niederle and Vesterlund 2007, Bandiera et al. 2013, Bracha and Fershtman 2013, Babcock et al. 2015). Lastly, our research aligns with a third strand, which investigates strategic interactions in longitudinal games with time-delayed actions and payments (e.g., Kim 2023).

The paper is organized as follows. Section I presents an overview of the experimental design and procedures. The model and the theoretical predictions are presented in Section II. In Section III we report the main results. In Section IV we analyze the Cooperative-obs and Competitive-obs treatments. Section V summarizes the results and concludes.

I. EXPERIMENTAL DESIGN

At the start of the study, all workers were informed that the study required them to log in and work remotely for three consecutive days. Their earnings depended on the number of transcription tasks they completed over the three days. Each transcription task required them to type a sequence of six randomly chosen letters correctly. Two days after completing their Day 3 tasks, the workers obtained their payment electronically through the MTurk platform.[3] Studying the intertemporal allocation of work over three days is crucial to our research question about procrastination. Hence, to discourage attrition and encourage participation on all three days, workers who participated on all three days were paid a \$4.00 log-in bonus. Failing to participate on any day disqualified them from the rest of the study, and they forfeited the \$4.00.

Figure 1 displays the experimental interface for the typing tasks. The interface contains the basic instructions and displays the minimum number of tasks that workers committed to type each day to be eligible for the \$4.00 log-in bonus (Figure 9 in Appendix F) and their plans.[4]

I.1. **Work Treatments.**

Workers were randomly assigned to one out of five treatments (three baseline treatments and two more derivative treatments that added observability). In the baseline *Cooperative* and *Competitive* treatments, they were randomly paired with another subject. In the Cooperative treatment, workers in a team of

[3]We needed at least a day to verify the group payments and pay all the subjects.

[4]The instructions can be found in Appendix D.

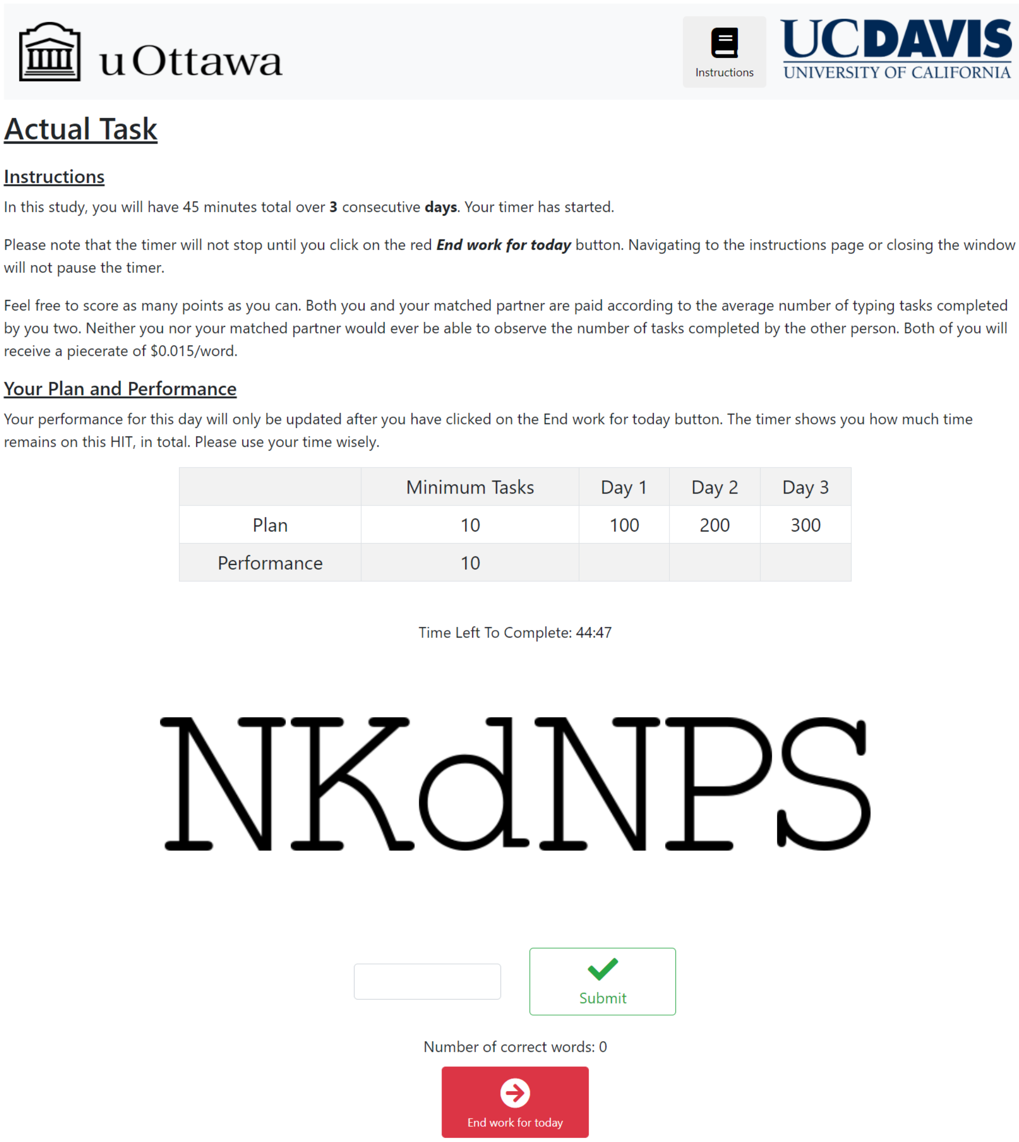


FIGURE 1. *Typing Task Interface*

two were paid equally based on their average performance and a piece rate p: Team members who completed e_1 and e_2 tasks respectively got $p(e_1 + e_2)/2$ each. Next, in the Competitive treatment, workers in a team of two got a minimum guaranteed piece rate of $2p/3$, but the worker completing more tasks among the pair earned an additional bonus payment $p(e_1 + e_2)/3$ that depended on the aggregate output. In the baseline team treatments, workers *could not observe* their partner's effort from past days. As a control, we ran an individualistic (*Solo*) treatment, where workers were paid based on their own effort: they get pe where p is the piece rate and e is the total number of typing tasks they completed. As illustrated in Figure 2, the payment schemes were designed such that all three treatments paid equally on average, thus making them comparable in terms of total worker compensation.

Our competitive treatment has two unique features compared to other experimental studies of competitive incentives: Firstly, we implement a bonus payment that increases with the total value generated by the team members, similar to how firms allocate performance bonuses by providing company stocks. In contrast, experimental studies of competitive incentives (e.g., Gneezy et al. 2003, Niederle and Vesterlund 2007) or studies of competitive institutions like contests or rank-order tournaments (see Dechenaux et al. 2015 for a review) usually offer a bonus that is exogenously fixed and independent of total effort.[5] We consciously avoided such a design as we believe not all competitive environments share that feature, and also because Bandiera et al. [2005] find that such competitive incentives can be self-defeating, leading workers to lower their effort under a relative-incentive scheme as they internalize the externality they impose on others.

Secondly, much of the experimental literature pays competition-winners disproportionately more than the losers. For example, the design in the well-replicated study by Niederle and Vesterlund [2007] pays competition-winners \$2 per correct problem solved and pays losers nothing. Similarly, designs by Bracha and Fershtman [2013] and Gneezy et al. [2003] pay the winner 6 times that of the loser. To give competition a fair shot, we removed such pay-disparity, which has been identified in the psychological literature as one of the environmental triggers of "choking under pressure" (Baumeister 1984, Baumeister and Showers 1986). We ensure that the losing party faces a smaller disadvantage by guaranteeing a substantial base piece rate of $2p/3$ to both participants.

Ex-post information available on payment date: In the baseline team treatments, partners could not observe each other's efforts during the three workdays. However, on their payment day, by receiving minimal feedback (information about their own payment), they could deduce their partner's effort[6]—completely in the Cooperative

[5]As an exception, Chen 2020 studies the effect of inter-team competition on team-members in a setting where resources endogenously generated by the losing team is transferred to the winning team. See Sheremeta 2018 for a review of the broader inter-team competition literature.

[6]Any team-based compensation scheme would always reveal information about the teammates' effort on the payment date.

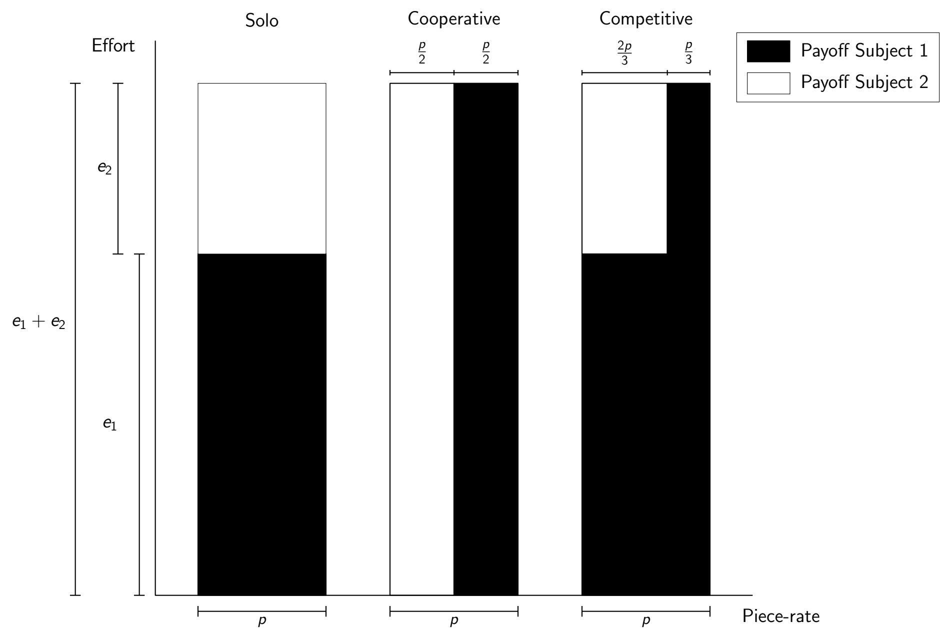


FIGURE 2. Payoff Equivalence for any piecerate p: Consider two workers who have completed e_1 and e_2 typing tasks respectively with $e_1 > e_2$. For each treatment, we plot their total payments as the areas in white and black. The areas of the three pillars, which represent the total payments made to the two workers under the three treatments, are identical and equal to $p(e_1 + e_2)$.

treatment and partially in the Competitive treatment.[7] In the Cooperative treatment, because the payments disclose the average effort, knowing one's own effort and payment (or average effort) is sufficient to infer the partner's effort. Consequently, both players can determine each other's effort, making this common knowledge. In the Competitive, failing to win the bonus reveals that the partner completed more tasks (partial information), whereas winning the bonus completely reveals the partner's effort based on the size of the bonus (full information). As with the Cooperative setup, this is also common knowledge.

Still, by not revealing partner's effort during Days 1-3, the two team treatments described thus far rule out workers changing their effort in retaliation to each other's effort *(Reaction channel* henceforth*)* or increasing work to influence their partner *(Signaling channel* henceforth*)*. To measure if feedback about the partner induces the *Reaction* and *Signaling* channels, we introduce two new variations to the baseline team treatments: the *Competitive-obs* and *Cooperative-obs* treatments. In the Competitive-obs and Cooperative-obs treatments, workers receive feedback on Days 2 and 3 about

[7]In practice, on their payment day, we also explicitly informed them about both their own and their partner's efforts each day. However, they were not aware of this at the beginning of the experiment.

the number of tasks completed by their partner on previous days.[8] Subjects are informed about this observability feature from the beginning of the study. For a detailed comparison of incentives and feedback across the treatments, please refer to Table 1.

We essentially *decompose* the informational stimulus present in teams into our baseline and *obs* treatments. The baseline isolates the effect of ex-post information, while the *obs* treatments isolates the effect of real-time feedback that can influence reactions as well as create anticipatory signaling effects.

TABLE 1. Experimental Treatments

Treatment	Payoff	Feedback About Partner	Number of Workers that Submitted Tasks	Number of Workers that Completed 3 Days	Average Payoff
Baseline Treatments					
Solo	pe	NA	125	103	\$ 9.16
Cooperative	$\frac{p}{2}(e_1+e_2)$ to both	No	103	91	\$ 9.60
Competitive ($e_1 > e_2$)	$\frac{2p}{3}e_1+\frac{p}{3}(e_1+e_2)$ and $\frac{2p}{3}e_2$	No	118	94	\$ 9.46
Obs Treatments					
Cooperative-obs	$\frac{p}{2}(e_1+e_2)$	Yes	122	91	\$ 9.60
Competitive-obs	$\frac{2p}{3}e_1+\frac{p}{3}(e_1+e_2)$ and $\frac{2p}{3}e_2$	Yes	129	90	\$ 10.63

Note: $p = \$0.015$ is the piece rate and e the total number of tasks they completed over three days. For the Cooperative and Competitive treatments, e_1 and e_2 refer to the total tasks completed by the best and worst performers, respectively. The payoffs include the log-in bonus.

I.2. **Experimental Procedures.** We recruited workers from Amazon Mechanical Turk (MTurk). The data-collection period spanned from June 2020 to January 2021. Participants accessed the study through the MTurk interface and were then redirected to the GUI designed on oTree (see Chen et al., 2016).[9] MTurk offers several advantages over standard university subject pools for our research on productivity and retention.

Firstly, it enabled us to recruit relatively large subject pools (just under 600 workers) while preventing direct communication between them *despite* the multi-day strategic interaction (MTurk workers are anonymous and spread all over the country,

[8]In Figure 11 within Appendix F, we provide a screenshot describing the feedback received by the participants in both the Competitive-obs and Cooperative-obs treatments.
[9]Table 6 in Appendix A summarizes the sample by key demographic characteristics.

unlike university students). This anonymity was crucial for maintaining the integrity of the experimental protocol, as any communication between paired partners outside the designated experimental period could compromise the comparison between the observability and baseline treatments.[10] Secondly, online workers, unlike students, would not have any future interactions with the researchers (us). As a result, they are less likely to feel obligated to keep participating in the tasks if they find it unrewarding or uninteresting. This helps us compare attrition under different incentives.

Only US workers who are authenticated via their social security number were allowed to participate in our study. To ensure that workers read and comprehended the instructions, they were required to successfully complete a series of comprehension checks that thoroughly assessed their understanding of the incentives and the intertemporal aspects of the task.[11] Out of the 1,297 workers who were directed to the oTree interface, 99 dropped out prior to reading the instructions, 59 dropped out before completing the quiz, and we excluded 526 participants who answered the majority of the quiz questions incorrectly. We also used attention checks hidden in the middle of the tasks to identify and remove inattentive workers and bots. The final sample consists of five treatments with 597 participants who passed the quiz and submitted at least one task. Out of these, 469 completed the minimum required tasks on each of the three days.

II. Model and Hypotheses

In this theory section, we show that our experimental design enables comparison between Solo versus team treatments under dominant strategies/one-round elimination of strictly dominated strategies. These comparative results hold true even in scenarios where (1) agents make choices over multiple days, (2) tradeoffs between costs and benefits occur at different times, and (3) there exists potential dynamic inconsistency due to one's incorrect beliefs about one's own future preferences. In Subsection II.1, we present a formal model of intertemporal effort choice in the Solo task, building upon the work of O'Donoghue and Rabin [1999]. This model incorporates the features (1)-(3) described above.

II.1. **Solo incentives model.** Workers are required to work on days 1,2, and 3. If they complete e_1, e_2, e_3 tasks respectively on these three days, they are paid $p(e_1 + e_2 + e_3)$

[10]If students participated alongside fellow students over multiple days in a lab, it would be impossible to track or control their communication outside the lab.

[11]Figure 8 in Appendix E shows the questions asked to workers in the Solo treatment. They have to answer the majority of the questions correctly to be able to proceed.

on a later day $T+1>3$. We assume that the effort cost is modelled by a function $c(e)$ with $c'>0, c''>0$. Suppose n is the number of tasks completed by them per minute. Workers cannot work for more than a total of 45 minutes over the three days, which translates to $45n = M$ total tasks.

$$e_1 + e_2 + e_3 \leq M$$

Assuming individuals have β-δ preferences, we define the dynamic problem recursively, starting from the last day. On the terminal day 3, Solo (s) workers decide how much they want to work on Day 3, denoted as $e_3^s(e_1, e_2)$, which is a function of their previous decisions e_1 and e_2 (state variables). The effort cost from previous days is sunk and therefore not included in the utility expression below. The monetary benefit from previous effort is also sunk, but is yet to be experienced, and hence we include it in the expression.

Day 3's problem (Solo):

$$\text{(II.1)} \quad e_3^s(e_1, e_2, \beta, p) \equiv \arg\max_{e_3} k\beta\delta^{T-2}p\sum_{t=1}^{3} e_t - c(e_3) \text{ subject to } 0 \leq e_3 \leq M - e_2 - e_1$$

where k is a constant that guarantees that money and effort-cost are expressed in the same units. Following O'Donoghue and Rabin [1999], in each non-terminal period, individuals choose their current effort believing that their future selves possess a present bias parameter denoted as $\hat{\beta} \geq \beta$.[12] This allows for *naive* time-inconsistent choice, as subjects might make an advanced choice under the wrong belief that their future selves have no present bias ($\hat{\beta} = 1$).

Thereafter, on Day 2, individuals decide $e_2^s(e_1)$ based on the already-spent time e_1 (state variable), and the belief that their Day 3 self would choose $\hat{e}_3 = e_3^s(e_1, e_2, \hat{\beta}, p)$ optimally for present bias $\hat{\beta}$.[13]

Day 2's problem (Solo):

$$e_2^s(e_1, \beta, p) \equiv \arg\max_{e_2} \beta\delta^{T-1}kp(e_1 + e_2 + \hat{e}_3) - c(e_2) - \beta\delta c(\hat{e}_3) \text{ subject to}$$

[12]Hence, when $\hat{\beta} = \beta$, it refers to individuals who are considered sophisticated, as they are fully aware of their future self-control issues and can accurately anticipate their future behavior. On the other hand, when $\hat{\beta} = 1$, it describes naive individuals who are entirely oblivious to their future self-control problems and believe that their future selves will behave in the same manner as their present selves desire. Intermediate values of $\hat{\beta}$ characterize partial naivete individuals.

[13]A naive subject mispredicts her future preferences and hence her future choice: $e_3^s(e_1, e_2, \hat{\beta}, p) \neq {e^s}_3(e_1, e_2, \beta, p)$.

$$(i) \quad 0 \le e_2 \le M - e_1$$
$$(ii) \quad \hat{e}_3 \equiv e_3^s(e_1, e_2, \hat{\beta}, p)$$

Similarly, on Day 1, the individual holds beliefs about Day 2 and Day 3 actions, $\hat{e}_2$ and $\hat{e}_3$, and solves:

$$e_1^s(\beta, p) \equiv \arg\max_{e_1} \beta\delta^T kp(e_1 + \hat{e}_2 + \hat{e}_3) - c(e_1) - \beta\delta c(\hat{e}_2) - \beta\delta^2 c(\hat{e}_3)$$

subject to

$$(i) \quad 0 \le e_1 \le M$$
$$(ii) \quad \hat{e}_2 \equiv e_2^s(e_1, \hat{\beta}, p)$$
$$(iii) \quad \hat{e}_3 \equiv e_3^s(e_1, \hat{e}_2^s, \hat{\beta}, p)$$

e_1, e_2, e_3 are real numbers that represent specific choices and $e_1^s(\cdot), e_2^s(\cdot), e_3^s(\cdot)$ are optimal choice functions that map state-variables (such as previous choices and present bias) to real numbers. $\hat{e}_1, \hat{e}_2, \hat{e}_3$ are defined recursively for the beliefs $\hat{\beta}$ about present bias. The superscript s denotes Solo incentives. Solving $e_1^s(\cdot), e_2^s(\cdot), e_3^s(\cdot)$ recursively yields real numbers $e_1^{s*}(p), e_2^{s*}(p), e_3^{s*}(p)$, which represent the optimal numerical solution under Solo incentives at piece rate p (note the additional asterisk). We can now state our main theoretical result:

Proposition 1. *Let a subject with preference parameters (β, δ) have $c(e) = \frac{e^{\gamma+1}}{\gamma+1}$ for $\gamma > 0$.*

(i) Early effort: e_1^{s*} is weakly increasing in β.

Next, assuming that on Day 1 and 2 she believes that she would exhaust all 45 minutes and assuming an interior optimum, we get

ii) On Day 1, she would complete $e_1^{s*} = \frac{Mf}{1+f}$ tasks, where $f^\gamma = \left(\beta\delta\frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}} + \beta\delta^2\frac{1}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}\right)$.

iii) On Day 2, she would complete $e_2^{s*} = \frac{(\beta\delta)^{\frac{1}{\gamma}}}{1+(\beta\delta)^{\frac{1}{\gamma}}}(M - {e^*}_1)$ tasks

iv) On Day 3, she would either complete $M - e_1^{s*} - e_2^{s*}$ tasks (and thus exhaust all 45 minutes) or stop short of it at $e_3^s = (k\beta\delta^2 p)^{1/\gamma}$ tasks at which point the marginal benefit and cost of working are exactly equal.

Part (i) establishes a monotonic relationship between procrastination and Day 1 effort: as β increases, the amount of work completed on Day 1 increases weakly as a lower amount of work is postponed to be completed later. Thus Day 1 effort, that is observable, can serve as a proxy for present bias (unobservable in our experiment).

Next, to derive a simple closed-form solution to the individual's problem, we assume an interior solution *and* that k is large enough such that the subject believes on Days 1 and 2 that she would exhaust all 45 minutes. This captures a typical naive worker who delays work for later days, expecting her future selves to exhaust all time. They can fall short of the goal when they get to the final day, if they underestimated the extent of their present bias ($\hat{\beta} > \beta$). This leads to simple closed form solutions for each day's effort which are also used later for our structural estimation.[14]

II.2. **Team incentives.** Under team incentives, the worker's beliefs about partner's actions matter. Let (e'_t) for $t = 1, 2, 3$ represent the tasks completed by the partner on Day $t = 1, 2, 3$ and let $e' = \sum_{t=1}^{3} e'_t$ denote its sum. Let b be the belief distribution over the three-tuple of partner's effort in the three periods. Next, we demonstrate that our design enables us to derive predictions regarding best responses that are *independent* of all such beliefs and that are robust under the assumption that players enjoy additional intrinsic motivation in teams.

Cooperative Treatment (Non-obs): We assume a simple model of *cooperation-based intrinsic motivation*, where the subject's final utility is the sum of her selfish utility and $\alpha \in [0, 1)$ times her team-mate's utility, where α is a constant utility parameter.[15] For any belief b,

$$e_3^{coop}(e_1, e_2, \beta, p) \equiv \arg\max_{e_3} k\beta\delta^{T-2} \left[\underbrace{\frac{p}{2}(\sum_{t=1}^{3} e_t + \boldsymbol{E_b} \sum_{\boldsymbol{t=1}}^{\boldsymbol{3}} \boldsymbol{e'_t})}_{\text{own pay}} + \alpha \underbrace{\frac{p}{2}(\sum_{t=1}^{3} e_t + \boldsymbol{E_b} \sum_{\boldsymbol{t=1}}^{\boldsymbol{3}} \boldsymbol{e'_t})}_{\text{partner's pay}} \right]$$
$$- \left(c(e_3) + \alpha \underbrace{\boldsymbol{E_b c(e'_3)}}_{\text{partner's cost}, \frac{d(\cdot)}{de_3} = 0} \right)$$
$$\text{subject to} \quad \underbrace{0 \leq e_3 \leq M - e_1 - e_2}_{\text{identical to Day 3 constraint from Solo}}$$

where all expectations E are taken w.r.t the belief b.

Remark 1. The terms containing partner's effort (see **bolded**) and effort from previous periods vanish from the first derivative of the Lagrangian, making beliefs irrelevant and resulting in a dominant strategy based on the marginal benefit. Specifically, the

[14]If one relaxes the interior optima assumption, part (iv) of our result still holds, but that proof is significantly longer has to account for multiple edge cases.
[15]α can also be interpreted as an altruism parameter.

marginal benefit is $(1+\alpha)/2$ times that of the Solo problem, and the marginal cost is identical to that of the Solo problem. Consequently, the Day 3 problem and optimal response for such a player are identical to those observed in the Solo treatment, with piece rate $p(1+\alpha)/2 \in [p/2, p)$, which is strictly smaller than p. In the absence of intrinsic motivation, the *effective piece rate reduces to* $p/2$. It can be easily demonstrated that the same applies to all other periods. Thus, under selfish preferences, a subject has a *dominant strategy* to exert an effort equal to the Solo-optimal at piece rate $p/2$, i.e, $e_t^{coop*}(p) = e_t^{s*}(p/2) \forall t$. Under an intrinsic motivation parameter $\alpha \in (0,1)$, she has a dominant strategy to exert an effort equal to the Solo-optimal at piece rate $p(1+\alpha)/2 \in [p/2, p)$, i.e, $e_t^{coop*}(p) = e_t^{s*}\left(p(1+\alpha)/2\right) \forall t$.[16] Therefore, we can compare the optimal strategy under Solo incentives with the dominant strategy under Cooperative incentives. We provide a formal result (Proposition 2) and proof in the Appendix.

Remark 2. (**Competitive Treatment**) Similarly, in the Competitive treatment, the player is guaranteed a piecerate of atleast $2p/3$. Each marginal unit of effort (i) increases her total payment by $2p/3$, (ii) increases the bonus she wins conditional on beating her competition, and (iii) weakly increases the chances of beating her competition. Thus, *irrespective of her beliefs* about partner's effort or the size of the bonus, she must be facing an effective piecerate of at least $2p/3$, which is still higher than $p/2$, the effective piecerate under Cooperative incentives. Thus we are led to the following hypotheses:

Hypotheses:

Hypothesis 1 (Productivity based on effective piecerates)**.** Under purely selfish preferences, workers in the Cooperative treatment should have a lower rate of exhausting 45 minutes, and a lower amount of time spent/work finished over the three days, compared to both the Solo and Competitive treatments.

Even under intrinsic motivation, since $\alpha < 1$, the effective piece rate under Solo *still* dominates that under the Cooperative treatment as $p(1+\alpha)/2 \in [p/2, p)$.[17] Cooperative and Competitive treatments can only be compared as long as

$$\begin{aligned} (1+\alpha)p/2 &< 2p/3 \\ \iff \alpha &< 1/3. \end{aligned}$$

[16]We provide a formal statement and proof in the Appendix.

[17]Comparing Solo and Competitive treatments *theoretically* requires artificial assumptions about the beliefs $F(\cdot)$ about one's partner, and we avoid taking this route.

This leads to our next hypothesis.

Hypothesis 2 (Productivity under Intrinsic Motivation)**.** i) Even under intrinsic motivation, workers in the Cooperative treatment should have a lower rate of exhausting 45 minutes, and a lower amount of time spent/work finished over the three days, compared to the Solo treatment.
ii) Under moderate intrinsic motivation ($\alpha < 1/3$), the Cooperative treatment would also be similarly dominated by the Competitive treatment.

Finally, if team incentives induce workers to procrastinate less, then teams would complete more tasks on Day 1, despite having lower piecerates.

Hypothesis 3 (Lower Procrastination in teams)**.** Ceteris paribus, workers complete more tasks on Day 1 in the two team treatments than in the Solo treatment.

II.3. **Observability treatments.** The observability condition in the Cooperative and Competitive treatments enables workers to observe their partners' effort levels on previous days and adjust their own behavior accordingly. We investigate two channels that separate the observability and non-observability conditions.

II.3.1. *Reaction Channel.* Observability implies that workers can know if their partner is working more than them, and then they can react to that information by increasing or decreasing their own effort- we call this the *Reaction channel*. Knowing that one has fallen behind in Competition can have one of two effects on the lagging worker. Either she can lose motivation, reduce her effort, and even drop out of the competition (*disillusionment effect*). Or she might increase her effort further to overcome the competitor (*resilience effect*). We choose the former as the null hypothesis and state it below. In the Cooperative treatment, reciprocity [Charness and Rabin, 2002] suggests that, ceteris paribus, workers should increase their effort if they observe their partners working harder.

Hypothesis 4 (Reaction Channel)**.** If a worker observes that their partner is working harder, it should decrease their effort under Competitive incentives and increase their effort under Cooperative incentives.

II.3.2. *Signaling Channel.* Apart from the Reaction channel, another distinction between the main and observability treatments is the scope of costly signaling through Day 1 effort. In the Competitive-obs treatment, for instance, they could choose to work harder on Day 1 to present themselves as formidable opponents, thereby dissuading their partner from competing on Day 2 onwards. We refer to this phenomenon as the *signaling hypothesis*. It also offers a clear directional prediction for cooperative incentives: workers are expected to exert more effort on Day 1 to demonstrate their virtue/ type and encourage their partner to reciprocate. Thus, the signaling effect, when present, would counteract present-biased tendencies of postponing work for later.

Hypothesis 5 (Signaling Channel)**.** Compared to the respective non-observability treatments, Day 1 effort should increase under Competitive and Cooperative incentives, when partner's effort becomes observable.

III. Results

In the first part of this section, we present the results for overall effort and time spent in the three treatments where observability was turned off. Our main goal is to analyze effort patterns over time, especially focusing on the potential impact of team treatments in reducing present bias. For this reduced-form analysis,we exclude all participants who failed the quiz or failed to engage with the multi-day feature of the study by exhausting all their time by Day 1 or 2. We still include subjects who followed all the instructions but failed to show up on a later day and thus got disqualified later on: our results are also robust to removing them from the data instead. A side-by-side comparison of attrition rates across treatments can be found in Appendix B. To disentangle how teams affect intrinsic motivation as well as procrastination, we perform a structural analysis of our hypotheses in subsection III.2. In Section IV, we discuss how feedback received on Days 2 in 3 (in the observability treatments) affects behavior.

III.1. **Productivity.** We begin our analysis by examining *total tasks completed* in columns [1]-[2] of Table 2. The odd-numbered columns ([1], [3], [5], [7]) measure the primary treatment effects, while the even-numbered columns offer robustness checks with the inclusion of demographic controls.[18] Compared to the Solo treatment, workers in the Cooperative and Competitive treatments completed approximately 65 and 35 additional tasks, but only the former difference is statistically significant. Therefore, while both team treatments result in greater productivity, the effect of the Cooperative treatment is not only more pronounced but also statistically significant.

[18] The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more).

The strong performance of the Cooperative treatment leads us to reject hypotheses 1 and 2.

III.1.1. *Importance of Day 1 productivity:* Is there an a intertemporal pattern that explains when and how teams work more? In columns [3] and [5] of Table 2, we compare the productivity on the first day (D1) and the last two days (D2 + D3), across the treatments. We find that both team treatments significantly and almost equally increased the amount of work completed on Day 1 (hypothesis 3).

The Day 1 increase in effort accounts for 36.4% (23.54/64.52) of the total productivity boost under Cooperative incentives, and 64.6% (22.20/35.44) of the total productivity increase under Competitive incentives. Even the 36.4% is quite high a fraction when compared to the percentage of total tasks completed on Day 1 by Cooperative workers on average: 107/273=26.8%. In contrast, in the Competitive treatment nearly two-thirds of the overall increase in total productivity can be attributed to the productivity gain on Day 1.

The Cooperative treatment also significantly increased the total productivity from Days 2 and 3 by 45.05 units, as indicated in column [5]. In contrast, the Competitive treatment had only a minor and insignificant effect on productivity during Day 2 and 3.

III.1.2. *Exhausting 45 minutes:* In columns [7]-[8], we define a categorical variable, *spent 45 minutes*, for subjects who exhausted all 45 minutes of work. To address any potential rounding issues, we classify participants who worked for strictly more than 44 minutes as having *spent 45 minutes*. All our results are robust to changing the cutoff to 44.5 minutes. The extra effort spent by team workers also show up in this measure: Both Cooperative and Competitive workers were significantly more likely to utilize the full allocated time. In case of the Competitive treatment, this difference can be attributed, to a large extent, to the higher productivity on Day 1.

III.2. **Structural estimation.** There are two potential mechanisms that explain why team incentives lead to better outcomes compared to Solo incentives. The first suggests that "team incentives have a higher motivational impact than the implied piece rates". This aligns well with the intertemporal pattern observed in the Cooperative treatment, especially with how effort increases on both Day 1, and on Days 2-3. *Intrinsic motivation* [Ryan and Deci, 2000], potentially driven by social preferences

Table 2. Regression Outcomes from Baseline treatments

	Dependent variable							
	Total No. of Tasks		Tasks on D1		Tasks on D2 + D3		1=Spent 45 mins 0=Otherwise	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Cooperative	64.520***	70.247***	23.540**	24.015**	45.049**	49.440***	0.184***	0.178***
	(23.34)	(23.15)	(9.44)	(9.67)	(18.72)	(18.55)	(0.07)	(0.07)
Competitive	35.444	38.841	22.206**	23.754***	13.699	15.622	0.112*	0.103
	(25.46)	(25.49)	(8.92)	(8.92)	(19.98)	(20.10)	(0.07)	(0.07)
Female		6.295		−20.810**		20.681		0.026
		(21.61)		(8.64)		(16.83)		(0.06)
Some College		22.490		1.334		20.521		0.082
		(32.03)		(15.57)		(23.84)		(0.10)
>Bachelor's Degree		33.001		-1.838		34.716		0.181*
		(30.26)		(15.71)		(21.86)		(0.09)
25-30 years old		−104.925***		−31.826**		−70.156**		-0.118
		(39.63)		(14.24)		(29.72)		(0.10)
31-40 years old		−105.403***		-22.067		−75.848***		-0.140
		(37.93)		(13.53)		(28.45)		(0.09)
41-50 years old		−127.717***		−34.056**		−90.900***		-0.047
		(40.18)		(13.90)		(30.10)		(0.10)
51-64 years old		−162.015***		−39.524**		−121.621***		-0.055
		(40.86)		(16.07)		(30.93)		(0.11)
>65 years old		−259.790***		−66.295***		−190.897***		-0.280
		(63.82)		(17.87)		(48.85)		(0.29)
Constant (Solo)	309.356***	381.985***	83.559***	123.795***	225.797***	258.795***	0.517***	0.465***
	(17.11)	(43.05)	(5.44)	(15.52)	(13.79)	(31.94)	(0.05)	(0.12)
Observations	320	320	330	330	320	320	320	320
	p values from F tests							
Cooperative vs Competitive	0.24	0.20	0.90	0.98	0.10	0.08	0.28	0.26

Note: Standard errors, which are robust, are presented in parentheses. We have omitted the dummy variable for 'Solo', meaning the constant reflects the baseline measure for Solo. All models use OLS regressions. The sample comprises workers who both completed the quiz and worked on the tasks. Those who exhausted 45 minutes on either the first or second day are not included. In columns [3] and [4], we also include 10 additional workers from the Cooperative and Competitive treatments who submitted tasks on Day 1 but could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days). The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

[Charness and Rabin, 2002], can complement the extrinsic monetary motivation provided by employers.[19] Similarly, even though a portion of the earnings in the Competitive treatment is at risk due to competition, workers' performance under this treatment was no worse than with Solo incentives. This supports the hypothesis that competitive incentives might also motivate beyond just the implied piece rates, as participants derive additional utility from winning or simply being competitive.

Second, consistent with the importance of Day 1 productivity reported in Table 2 (see Hypothesis 3), we suggest as a novel mechanism: "team incentives mitigate present-biased behavior." Present bias has been described as the inability to empathize with one's future selves [Loewenstein, 2005], and is generally considered an immutable preference parameter. There still exists empirical evidence [Halevy, 2015] as well as theoretical modeling [Chakraborty, 2021, Noor and Takeoka, 2022] demonstrating situations where present bias varies depending on the circumstances and stakes involved.

For example, Noor and Takeoka [2022] explore how individuals optimize the cognitive cost of empathizing with their future selves based on their specific circumstances. Building upon this insight, our second proposed mechanism suggests that when individuals encounter team incentives, they re-optimize the empathy they extend towards their future selves, which reduces present bias and subsequently enhances productivity. The observed productivity boost aligns well with these insights. To assess and analyze the two mechanisms as well as their combined effects, we conduct a structural estimation exercise.

III.2.1. *Auxiliary parameters that are not estimated:* To identify the parameters of interest, we need to assign values to two auxiliary parameters.

- Following Augenblick et al. [2015], we assume a parametric effort-cost function $c(e) = \frac{e^{\gamma+1}}{\gamma+1}$ for completing e tasks. We adapt their estimated cost parameter $\gamma = 0.589$.
- As the daily discount factor, we adopt Augenblick et al. [2015]'s estimated $\delta = (.999)^{1/7}$. Assuming $\delta = 1$ would result in identical estimates.
- We assume $\hat{\beta} = 1$, i.e, complete näiveté.

III.2.2. *Model:* We assume that each day, subjects draw an independent productivity shock through their present bias, denoted as β_t, which follows a Lognormal distribution with parameters μ and σ. This productivity shock plays a crucial role in

[19]Guilt aversion [Charness and Dufwenberg, 2006, Battigalli and Dufwenberg, 2007] or a strategic notion of desert [Gill and Stone, 2015] could also play a role in such intrinsic motivation.

determining their intertemporal tradeoff. We use the closed-form expressions derived in Proposition 1 for each day's effort. Thus on Day 1, workers complete $e_1 = \frac{Mf}{1+f}$ tasks, where $f^\gamma = \beta_1 \delta \frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}} + \beta_1 \delta^2 \frac{1}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}$, and M is the total number of tasks they would finish in 45 minutes of work. β_1 is the realized present bias for Day 1, and $\hat{\beta} = 1$. Next, conditional on working e_1 minutes on Day 1, on Day 2 they would decide to work for $e_2 = \frac{(\beta_2\delta)^{\frac{1}{\gamma}}}{1+(\beta_2\delta)^{\frac{1}{\gamma}}}(M - e_1)$ minutes, where β_2 is the realized present bias for Day 2. Finally, on Day 3, workers would either finish $e_3 = M - e_1 - e_2$ tasks or stop short of it at $e_3 = e_3^{max}$ tasks at which point the marginal benefit and cost of working are exactly equal. Note that the value of $e_3^{max} = (k\beta_3\delta^2 p_i)^{1/\gamma}$ is jointly determined by the realization of β_3 and the effective piece rate $p_i \in \{p_s, p_{coop}, p_{comp}\}$ for the specific treatment.[20] By imposing the values $\delta = (.999)^{1/7}$, $\hat{\beta} = 1$ and $\gamma = 0.589$ into the expressions, we can simplify and express e_t as simple functions of the realized value β_t and observed past effort.

For those who actually completed all 45 minutes, we observe M and use it directly. For those who do not, we impute M as the product of total time (45 minutes) and the average number of tasks they completed per minute over three days.

Since both k and p_i only appear in the e_3^{max} expression and appear multiplicatively, the point estimate of k depends on the unit (dollars, cents, euros etc.) in which p is expressed. To facilitate comparisons, we *normalize* the piece rate p_s to 1 for the Solo treatment. Thus, $p_{coop} = 0.5$ in the Cooperative treatment, half of that in the Solo treatment. As for the Competitive treatment, we equate the effective piece rate to the minimal piecerate $p_{comp} = 2/3$.[21]

III.2.3. *Specifications:* To enable team incentives to perform beyond the effective piecerates, we use and compare two different specifications: In specification [1], we introduce efficiency factors A_{coop} and A_{comp} to multiply the implied piece rates, which quantify the effectiveness of team incentives (and the intrinsic motivation therein) beyond their implied marginal piece rate. In particular, for the Cooperative treatment, we set $p_{coop} = 0.5A_{coop}$, which means $e_3^{max}|_{coop} = (k\beta_3\delta^2 0.5A_{coop})^{1/\gamma}$. Similarly, for the Competitive treatment, we use $p = 2/3 A_{comp}$ and hence $e_3^{max}|_{comp} = (k\beta_3\delta^2 2/3 A_{comp})^{1/\gamma}$. If $A_{coop} = A_{comp} = 1$, it suggests that team incentives do not provide any additional motivation beyond their implied marginal piece rate. On the other hand, if A_{coop} and

[20]The i in p_i stands for incentive.

[21]If we used $p_{comp} = 1$ instead, it would still only change the estimate of A_{comp} defined below, leaving the other parameters unchanged. $p_{comp} = 2/3$ is simply a normalization which shuld be used when interpreting A_{comp}.

A_{comp} are greater than 1, it indicates that "team incentives motivate more than the implied piece rates".

In specification [2], we remove efficiency factors by forcing $A_{coop} = A_{comp} = 1$, but allow the distribution of present bias to be $\beta_s \sim Lognormal(\mu_s, \sigma)$ for Solo and $\beta_{team} \sim Lognormal(\mu_{team}, \sigma)$ for both team treatments. One can interpret this as if "team incentives decrease present bias in temporal preferences". If we estimate $\mu_{team} = \mu_s$, then it suggests that team incentives do not provide any additional benefits. However, if the estimated means of the present bias distribution differ between treatments, the treatment with a higher mean would induce its participants to exert more effort.

In specification [3], we allow for the joint hypothesis that teams not only motivate more than the implied piecerate but also reduce procrastination. Specification [4] is a robustness check for [3], where we allow Cooperative and Competitive treatments to have their individual distributions of present bias.

We perform a Maximum Likelihood exercise to determine the lognormal distribution and the parameters α, σ, and k that jointly maximize the likelihood of the observed data. We include all subjects who submitted work on all three days.[22]

To explain the identification process, consider the following: For the fixed values of γ and δ, each observed time of work (e_1, e_2) for an individual j, within a specific treatment i, maps to a unique realization of β_t from the distribution. Consequently, the complete data on effort choices for any treatment can be fitted to the distribution $\beta_t \sim Lognormal(\mu, \sigma)$. Furthermore, the variation in e_3 between the Solo and team treatments helps in identifying the values of A_{coop} and A_{comp} respectively. For example, compare to Solo, workers in the Cooperative treatments worked significantly more on Day 3 and were 19% more likely to exhaust their M: this would be attributed jointly to a lower procrastination and to $A_{coop} > 1$.

III.2.4. *Results.* We report the results in Table 3a. In specification [1], the median[23] subject exhibits present bias with a value of $\beta = \exp(-.2) = 0.75$. Cooperative incentives appear to be more motivating, as evidenced by $A_{coop} > A_{comp} > 1$. This observation aligns with conclusions derived from our earlier reduced-form results. It

[22]The presence of attritioned workers who work zero minutes implies that $\beta_t = 0$ within the structural model, which is a zero measure draw from any lognormal distribution. If we run the maximum likelihood estimation (MLE) program including all attritioned subjects, the program still evaluates the same parameters and standard errors. However, the maximized likelihood significantly decreases due to the presence of zeros in the data.

[23]Since the Normal distribution is symmetric and unimodal, its mean and median are equal.

further suggests that both team treatments outperform what is predicted solely by the marginal piece rates. Moreover, A_{coop} is found to be significantly higher than 2, implying that intrinsic motivation more than compensates for the lower implied piece rate.

In specification [2], our results suggest that the median subject exhibits present bias, with values of $\beta = 0.63$ in the Solo treatment and $\beta = 0.81$ in the team treatments. Both of these parameters are statistically less than 1. Furthermore, a comparison shows that $\mu_s < \mu_{team}$, and this difference is statistically significant at the 1% level. This implies that teams enhance productivity by mitigating present bias, a finding that aligns with the productivity insights from our earlier reduced-form analysis.

Finally, specification [3] combines the features of models [1] and [2], incorporating both efficiency factors and different distributions of present bias. Thus, [3] encompasses both [1] and [2], and our tests indicate that this is the best model. Using a Likelihood Ratio test of nested models, we reject the hypothesis that [2] is a better model than [3] ($p < .01$), and that [1] is a better model than [3] ($p = .07$). Similarly, the Akaike Information Criterion (AIC) favors [3] over the other two specifications. This implies that the conjunction hypothesis—"team incentives decrease present-biased behavior and motivate beyond the implied piece rate"—fits our data better than the explanations of "team incentives decrease present-biased behavior" or "team incentives motivate beyond the implied piece rates". Thus, the reduction in present bias induced by teams explains part of the efficiency of team incentives.

When allowing for distinct present bias distributions across the two team incentives, as demonstrated in specification [4], the increase in explanatory power is barely noticeable, as evidenced by the loglikelihood. Therefore, model selection tests do not favor specification [4].

IV. Behavior under feedback

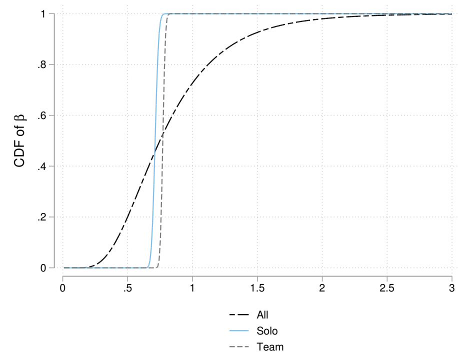
IV.1. **Reaction Channel.** To investigate the Reaction channel (Hypothesis 4), we exploit the random assignment of worker i's partner j, and by extension, the randomness of partner j's effort on Day 1. We regress i's effort on Day 2 ($e_{i,2}$) on j's effort on Day 1 ($e_{j,1}$) for both the Cooperative and Competitive treatments. In this context, i's Day 1 effort ($e_{i,1}$) is used as a proxy for worker i's inherent ability or attributes. We also model the reaction to a partner's past effort depending on whether it is perceived as "favorable" or "unfavorable" news by interacting $e_{j,1}$ with a dummy variable, $1_{e_{i,1}>e_{j,1}}$, which equals one if i's Day 1 effort exceeds that of j.

Table 3. Structural Analysis

(a) Structural estimates

	[1]	[2]	[3]	[4]
μ_s		$-.45^{***}$	$-.34^{***}$	-0.34^{***}
		.03	.03	(0.03)
μ_{team}		$-.20^{***}$	$-.26^{***}$	
		.02	.02	
μ_{coop}				-0.27^{***}
				(0.03)
μ_{comp}				-0.26^{***}
				(0.03)
μ	-0.29^{***}			
	(0.001)			
σ	0.48^{***}	0.50^{***}	0.47^{***}	0.47^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
k	23.30^{***}	35.83^{***}	24.34^{***}	24.33^{***}
	(1.10)	(1.32)	(1.41)	(0.87)
α_{coop}	2.50^{***}	1	2.31^{***}	2.33^{***}
	(0.18)		(0.19)	(0.16)
α_{comp}	1.53^{***}	1	1.41^{***}	1.40^{***}
	(0.10)		(0.11)	(0.09)
Obs	288	288	288	288
LL	-630.40	-675.78	-628.78	-628.75
AIC	1270.80	1359.561	1269.56	1271.50

(b) Estimated CDF of β



Note: The analysis includes all subjects who worked on all three days, and thus have no missing effort data. In the left panel, we estimate the distribution of β and the efficiency of team-incentives under specifications [1]-[3]. Standard errors are reported in parentheses. In the right panel, we plot the estimated log-normal distribution of β: The common distribution for all treatments from [1] is plotted under All, whereas the two separate distributions from [3] are plotted under Solo or Team incentives.

In scenarios where feedback is absent, like in the non-observable Cooperative and Competitive treatments, neither $e_{j,1}$ nor the dummy $1_{e_{i,1}>e_{j,1}}$ should influence $e_{i,2}$. However, our placebo tests in Columns [1]-[2] of Table 4 show an unexpected effect in the non-observable Competitive treatment, where Day 2 effort appears influenced by the partner's Day 1 effort, albeit only significant at the 10% confidence level. We interpret this as potentially occurring by chance.

We reject the disillusionment effect (Hypothesis 4) in favor of resilience in the Competitive-obs treatment: As $e_{j,1}$ rises, i boosts their effort regardless of whether competitor j is ahead or behind. This effect amplifies when i knows they're leading. Furthermore, with other factors controlled for, Day 1's effort $e_{i,1}$ is no longer a significant predictor of $e_{i,2}$.

In the Cooperative-obs treatment, an increase in $e_{j,1}$ usually spurs an uptick in $e_{i,2}$. However, this relationship becomes statistically insignificant once demographic controls are added, as shown in Table 10in Appendix I. The observed increment under cooperative incentives is notably less pronounced than under competitive ones.

IV.2. **Signaling Channel.** The signaling effect, when present, would counteract present-biased tendencies of postponing work for later. Contrary to our expectations (Hypothesis 5), we find little aggregate evidence of the signaling hypothesis in both competitive and cooperative incentives. On average, the first-day effort is 107.09 under Cooperative and 105.73 under Cooperative-obs. For the Competitive treatment, the average first-day effort is 105.77 and 106.04 under Comp-obs. When performing rank sum tests, where the null hypothesis is that the effort on Day 1 is the same under observability and non-observability, we fail to reject the hypothesis for both the cooperative (p-value = 0.7915) and the competitive (p-value = 0.8612) treatments.[24] To understand the absence of evidence for signaling, we conducted supplementary experiments using participants from Amazon Mechanical Turk (MTurk). Participants (referred to as "predictors") made incentivized guesses regarding the performance of earlier "worker" participants in both Cooperative and Competitive settings. These experiments aimed to explore five potential reasons for the lack of collective evidence supporting the signaling hypothesis. The potential explanations examined included heterogeneous beliefs, flat beliefs, unfavorable beliefs in competition, a preference to gather more information before acting, and a lack of strategic thinking. Our findings

[24]The absence of an effect under either incentive suggests that other factors beyond self-canceling heterogeneous effects may be influencing the outcome.

Table 4. Effect of Partner's past Effort

	Dependent variable: Number of Tasks on Day 2			
	Cooperative	Competitive	Cooperative-obs	Competitive-obs
	[1]	[2]	[3]	[4]
$e_{j,1}$	0.074	-0.178^{*}	0.126^{*}	0.340^{**}
	(0.09)	(0.09)	(0.07)	(0.14)
$e_{i,1}$	0.219^{*}	0.388^{**}	0.543^{***}	0.204
	(0.11)	(0.17)	(0.10)	(0.17)
$(1_{e_{i,1}>e_{j,1}}) \times e_{j,1}$	0.050	-0.032	0.245^{**}	0.460^{**}
	(0.15)	(0.24)	(0.12)	(0.21)
Constant	91.789^{***}	102.493^{***}	52.089^{***}	64.416^{***}
	(12.81)	(17.25)	(13.52)	(17.24)
Observations	94	96	94	92

Note: Standard errors, reported in parentheses, are robust across all treatments. For treatments with observability, they are clustered at the team level. In our notation, $e_{j,1}$ denotes the number of tasks submitted by the partner on Day 1, while $e_{i,1}$ represents the worker's own effort on the same day. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. Workers that could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

suggest that workers lacked the strategic thinking required for the signaling channel to work. Details can be found in Appendix C.

IV.3. **Aggregate effects of observability.** In Table 5, we present a comprehensive comparison between the incentive schemes, emphasizing versions both with and without observability. Observability does not affect the total tasks completed or time spent in a statistically significant manner. As we analyze the distribution of effort across time, we find that present bias is completely immune to the additional observability made available in the *obs* treatments: The signaling effect does not increase Day 1 effort in either team treatment. But, the reaction effect has a net positive and significant effect in the Competitive treatment for Days 2 and 3. Thus, observability does enhance overall productivity.

The introduction of immediate feedback in *obs* treatments alters the productivity ranking of the incentive schemes. Without it, cooperative incentives consistently outperform others across all comparison metrics. However, when a partner's past effort becomes observable on Days 2 and 3, competitive incentives are more productive than cooperative incentives, although the difference is not statistically significant. This is detailed in Table 12 found in the Appendix I.

Table 5. Effect of observability on team treatments

	Dependent variable			
	Total No. of Tasks	Tasks on D1	Tasks on D2 + D3	1=Spent 45 mins 0=Otherwise
	[1]	[2]	[3]	[4]
Cooperative	373.876***	107.099***	270.845***	0.701***
	(15.88)	(7.71)	(12.66)	(0.05)
Competitive	344.800***	105.766***	239.495***	0.629***
	(18.85)	(7.07)	(14.46)	(0.05)
Cooperative×obs	-24.235	-1.373	-24.506	-0.041
	(24.09)	(9.93)	(18.96)	(0.07)
Competitive×obs	41.181	0.277	39.486*	0.079
	(26.93)	(9.67)	(21.90)	(0.06)
Observations	414	442	414	414

Note: Robust standard errors are reported in parentheses. All specifications are OLS regressions. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. Workers that could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are excluded too. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

V. Concluding remarks

In this study we compare effort allocation in under solo and team-based incentives. Our evidence suggests that team incentives enhance overall productivity by providing intrinsic motivation *and* by reducing procrastination. Furthermore, we investigate additional mechanisms arising from team-based incentives, including the Reaction effect and the Signaling effect. We find compelling evidence of the Reaction effect, showing that workers increase their effort in response to their teammate's effort under both competitive and cooperative scenarios. Surprisingly, however, we do not find any evidence of the Signaling effect, suggesting that workers do not leverage their opponents' responses for their advantage. Our study contributes to multiple experimental literatures, touching upon intertemporal motivation, social compensation schemes, and strategic interactions in longitudinal games. Overall, our research sheds light on the potential of team-based incentives to enhance productivity, decrease procrastination, and foster intrinsic motivation among individuals.

References

Dan Ariely and Klaus Wertenbroch. Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science*, 13(3):219–224, 2002.

Ned Augenblick, Muriel Niederle, and Charles Sprenger. Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics*, 130(3): 1067–1115, 2015.

Philip Babcock, Kelly Bedard, Gary Charness, John Hartman, and Heather Royer. Letting down the team? social effects of team incentives. *Journal of the European Economic Association*, 13(5):841–870, 2015.

Timothy Baldwin, William Bommer, and Robert Rubin. *Developing management skills: What great managers know and do*. McGraw-Hill/Irwin, 2008.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 120(3):917–962, 2005.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social incentives in the workplace. *The review of economic studies*, 77(2):417–458, 2010.

Oriana Bandiera, Iwan Barankay, and Imran Rasul. Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, 11(5): 1079–1114, 2013.

Christopher M Barnes, John R Hollenbeck, Dustin K Jundt, D Scott DeRue, and Stephen J Harmon. Mixing individual incentives and group incentives: Best of both worlds or social dilemma? *Journal of Management*, 37(6):1611–1635, 2011.

Pierpaolo Battigalli and Martin Dufwenberg. Guilt in games. *American Economic Review*, 97(2):170–176, 2007.

Roy F Baumeister. Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, 46(3):610, 1984.

Roy F Baumeister and Carolin J Showers. A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, 16(4):361–383, 1986.

Stefania Bortolotti, Giovanna Devetag, and Andreas Ortmann. Group incentives or individual incentives? a real-effort weak-link experiment. *Journal of Economic Psychology*, 56:60–73, 2016.

Anat Bracha and Chaim Fershtman. Competitive incentives: working harder or working smarter? *Management Science*, 59(4):771–781, 2013.

Anujit Chakraborty. Present bias. *Econometrica*, 89(4):1921–1961, 2021.

Gary Charness and Martin Dufwenberg. Promises and partnership. *Econometrica*, 74(6):1579–1601, 2006.

Gary Charness and Matthew Rabin. Understanding social preferences with simple tests. *The quarterly journal of economics*, 117(3):817–869, 2002.

Daniel L Chen, Martin Schonger, and Chris Wickens. otree–an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

Yi-Yi Chen. Intergroup competition with an endogenously determined prize level. *Journal of Economic Behavior & Organization*, 178:759–776, 2020.

Jonathan Cohen, Keith Marzilli Ericson, David Laibson, and John Myles White. Measuring time preferences. *Journal of Economic Literature*, 58(2):299–347, 2020.

Maribeth Coller and Melonie B Williams. Eliciting individual discount rates. *Experimental Economics*, 2:107–127, 1999.

Emmanuel Dechenaux, Dan Kovenock, and Roman M Sheremeta. A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669, 2015.

Shane Frederick, George Loewenstein, and Ted O'donoghue. Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2):351–401, 2002.

David Gill and Rebecca Stone. Desert and inequity aversion in teams. *Journal of Public Economics*, 123:42–54, 2015.

Uri Gneezy and John A List. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5):1365–1384, 2006.

Uri Gneezy, Muriel Niederle, and Aldo Rustichini. Performance in competitive environments: Gender differences. *The quarterly journal of economics*, 118(3):1049–1074, 2003.

Noah J Goldstein, Robert B Cialdini, and Vladas Griskevicius. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research*, 35(3):472–482, 2008.

Yoram Halevy. Time consistency: Stationarity and time invariance. *Econometrica*, 83(1):335–352, 2015.

Alex Imas. Working for the warm glow on the benefits and limits of prosocial incentives. *Journal of Public Economics*, 114:14–18, 2014.

Jeongbin Kim. The effects of time preferences on cooperation: Experimental evidence from infinitely repeated games. *American Economic Journal: Microeconomics*, 15 (1):618–637, 2023.

John A List and Rohen Shah. The impact of team incentives on performance in graduate school: Evidence from two pilot rcts. Technical report, National Bureau of Economic Research, 2022.

George Loewenstein. Hot-cold empathy gaps and medical decision making. *Health psychology*, 24(4S):S49, 2005.

Muriel Niederle and Lise Vesterlund. Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101, 2007.

Jawwad Noor and Norio Takeoka. Optimal discounting. *Econometrica*, 90(2):585–623, 2022.

Ted O'Donoghue and Matthew Rabin. Doing it now or later. *American Economic Review*, 89(1):103–124, March 1999. doi: 10.1257/aer.89.1.103. URL http://www.aeaweb.org/articles?id=10.1257/aer.89.1.103.

Glenn Parker, Jerry McAdams, and David Zielinski. Rewarding teams: Lessons from the trenches. *Team Performance Management: An International Journal*, 2000.

Mitesh S Patel, David A Asch, Roy Rosin, Dylan S Small, Scarlett L Bellamy, Kimberly Eberbach, Karen J Walters, Nancy Haff, Samantha M Lee, Lisa Wesby, et al. Individual versus team-based financial incentives to increase physical activity: a randomized, controlled trial. *Journal of general internal medicine*, 31:746–754, 2016.

Emma Pearson, Harry Prapavessis, Christopher Higgins, Robert Petrella, Lauren White, and Marc Mitchell. Adding team-based financial incentives to the carrot rewards physical activity app increases daily step count on a population scale: a 24-week matched case control study. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1):1–10, 2020.

Richard M Ryan and Edward L Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1):54–67, 2000.

Roman M Sheremeta. Behavior in group contests: A review of experimental research. *Journal of Economic Surveys*, 32(3):683–704, 2018.

Appendix A. demographic information

Table 6. Mturk Sample

	Mean				
	Solo	Cooperative	Cooperative-obs	Competitive	Competitive-obs
Female	0.60	0.61	0.60	0.62	0.58
Education					
High School or Less	0.11	0.10	0.08	0.08	0.11
Some College	0.25	0.26	0.35	0.27	0.28
Bachelor's Degree or more	0.64	0.64	0.57	0.65	0.61
Age					
18-24 years old	0.10	0.14	0.11	0.11	0.14
25-30 years old	0.29	0.23	0.24	0.26	0.23
31-40 years old	0.34	0.32	0.34	0.33	0.31
41-50 years old	0.17	0.17	0.18	0.17	0.15
51-64 years old	0.10	0.13	0.11	0.11	0.13
Older than 65	0.01	0.02	0.02	0.02	0.04
Observations	210	198	262	248	280

Note: This table includes the 1424 workers that filled out the demographic questionnaire and for which the instructions were displayed.

Appendix B. Attrition

Due to the multi-day nature of our study, online workers who experience a lack of motivation or dissatisfaction with their compensation sometimes choose to leave before completing Day 2 or Day 3. In column [1] of Table 2, we observe that 14.9% of Solo workers and 12.1% of Competitive workers did not complete the full 3-day experiment. In contrast, the Cooperative treatment exhibited a significantly lower attrition rate of 7.1%. This means that workers in the Cooperative treatment were almost 50% less likely to drop out, compared to those in the Solo or Competitive treatments. This finding is also robust to adding demographic controls, as shown in column [2].

Table 7. Attrition

	Dependent variable: 1=did not log in for all 3 days; 0= Otherwise	
	[1]	[2]
Cooperative	-0.077^{*}	-0.087^{**}
	(0.04)	(0.04)
Competitive	-0.027	-0.038
	(0.05)	(0.05)
Constant (Solo)	0.149^{***}	0.028
	(0.03)	(0.05)
Controls	No	Yes

Note: Robust standard errors are reported in parentheses. All specifications are OLS regressions. Workers from the Cooperative and Competitive treatments who submitted tasks on Day 1 but could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B.1. **Attrition and Gender.** A prominent experimental literature (e.g., Niederle and Vesterlund [2007]) suggests that a higher proportion of men opt into competitive environments compared to women. Do women also exhibit a higher likelihood of dropping out from jobs that offer competitive incentives? Table 8 demonstrates that the answer is yes, but only when workers receive feedback about their partners. However, we do not find a significant gender pay gap under competition. When considering participation on Day 1 of Competition-obs, women earn approximately \$0.80 less than men. When considering completion of all three days, women earn even less, approximately \$0.50 less than men. This difference is partly explained by women being significantly more likely to discontinue their participation. In both cases, the gap is not statistically significant.

TABLE 8. Attrition and Gender

	Dependent variable: 1=did not log in for all 3 days; 0= Otherwise				
	Solo	Cooperative	Cooperative-obs	Competitive	Competitive-obs
Female	-0.075	-0.003	-0.047	-0.115	0.130**
	(0.07)	(0.06)	(0.07)	(0.08)	(0.06)
Constant	0.045	-0.068	-0.029	0.206	0.086
	(0.07)	(0.08)	(0.08)	(0.16)	(0.12)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	121	98	107	108	109

Note: Robust standard errors are reported in parentheses. All specifications are OLS regressions. Workers from the Cooperative and Competitive treatments who submitted tasks on Day 1 but could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). Standard errors are clustered at the group level for the observability treatments. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix C. Exploring the lack of signaling

To investigate the lack of evidence for signaling, we recruited a new group of participants, whom we will refer to as "predictors" to study their guesses. We make the identification assumption that subjects who are randomly assigned to the roles of predictor or worker hold similar beliefs about strangers, allowing us to study the beliefs of the original worker population. Half of the predictors were randomly assigned to prediction tasks about the Cooperative environment, while the other half were assigned to prediction tasks about the Competitive environment. We name the first treatment Coop-pred and the second one Comp-pred. Predictors were given the same instructions as in the respective treatment and were asked to complete the same comprehension quiz. However, instead of completing transcription tasks themselves, they were asked to make incentivized guesses about the performance of workers who had previously participated.[25]

In the Coop-pred treatment, we paired predictors with past participants and asked them to make predictions about workers who had taken part in the Cooperative and Cooperative-Obs treatments. Similarly, in the Comp-pred treatment, predictors made guesses about previous workers who had participated in the Competitive and Competitive-Obs treatments. Each predictor made incentivized guesses about the items shown in Table 9.[26] For each guess the participants were paid following this formula $\max\left\{\$0, \$0.50 - \left(\frac{\text{guess}-\text{actual value}}{100}\right)^2\right\}$. For the second part of the experiment, participants were asked to choose between three different actions in two 3×3 matrix games.[27]

We designed the prediction treatments to investigate five potential explanations for the lack of aggregate evidence supporting the signaling hypothesis channel:

[25] We recruited participants from the pool of workers on Amazon Mechanical Turk (MTurk). There were 98 participants in the Coop-pred treatment group and 99 participants in the Comp-pred treatment group.

[26] We only allow participants who passed the quiz to submit their guesses. Additionally, we remove predictors from the sample if at least one of their guesses was below the lower bound when provided information. For example, we remove a predictor if their guess for $e_B^{(i)}$ was lower than 20.

[27] The two matrices were identical. In the first game, participants were assigned the role of the row player, while in the second game, they were assigned the role of the column player. However, to simplify the experiment, we transposed the first matrix, so in the second game, participants also had the role of row players. At the end of the experiment, one of the two games was randomly selected for each predictor. If the first game was chosen, then participants were paired with a partner for whom the second game was selected, and vice versa. Each participant was then paid the payoff associated with the choice they made in the selected game and for the decision made by their partner in the other game. Appendix G shows the screenshots for the instructions.

Table 9. Guesses

Worker	Guess	Definition
$\emptyset$	$e_{\emptyset}^{(i)}$	Total effort without any information about the worker
l	$e_{l}^{(i)}$	Total effort of a worker that completed between 20 and 30 tasks on Day 1
$l´$	$e_{l'}^{(i)}$	Total effort exerted by worker $l´$s partner
h	$e_{h}^{(i)}$	Total effort of a worker that completed between 100 and 110 tasks on Day 1
$h´$	$e_{h'}^{(i)}$	Total effort exerted by worker $h´$s partner
ll	$e_{ll}^{(i)}$	Total effort of a worker that completed between 50 and 80 tasks on Days 1 and 2
hh	$e_{hh}^{(i)}$	Total effort of a worker that completed between 220 and 250 tasks on Days 1 and 2
obs	$e_{obs}^{(i)}$	Day´s 1 effort of worker that participated in the Obs treatment
$nobs$	$e_{nobs}^{(i)}$	Day´s 1 effort of worker that participated in the No-Obs treatment

Note: The superscript indicates that the guess is being made by predictor i. These guesses were compared with the actual effort of randomly selected workers who participated in either the Cooperative or Competitive treatments. The guesses were collected at the treatment level (Coop-Pred and Comp-Pred). For each prediction, participants were also requested to indicate their level of confidence.

H0) Heterogeneous beliefs: An equal number of workers believe that increasing their Day 1effort will increase or decrease their partner's total effort, resulting in a cancellation of effects in the aggregate data.

H1) Flat beliefs: Workers individually do not believe that increasing their Day 1 effort will have any impact on their partner's total effort in either the Cooperative or Competitive incentive treatments.

H2) Unfavorable beliefs in Competition: Workers in the competitive treatment believe that working harder on Day 1 will lead to increased effort from their partner, which would lower their chances of winning the bonus.

H3) Information Motive: Workers wish to wait until Day 2 before committing significant effort to have a more informed estimation of their partner's total effort over three days.

H4) Lack of Strategic Thinking: Workers lack the strategic reasoning or "theory of mind" required to think through the implications of their Day 1 effort.

We can directly test H0, H1 and H2 by comparing whether $e_{l'} = e_{h'}$. This test assesses whether predictors expect that the high and low effort of subject h or l on

Day 1 will have no influence on their partner's total effort. Sign tests reject $e^{(i)}_{l'} = e^{(i)}_{h'}$ in favor of $e^{(i)}_{l'} < e^{(i)}_{h'}$ for both treatments (p-value=0.000 for both).[28]

To directly test H3, we compare the self-reported confidence levels of predictors who make guesses with knowledge of Day 1 effort versus those with knowledge of (Day 1 + Day 2) effort. For example, we can compare the self-reported confidence levels of $c^{(i)}_{l}$ and $c^{(i)}_{ll}$ for the predictions $e^{(i)}_{l}$ and $e^{(i)}_{ll}$, respectively. If we find that $c^{(i)}_{ll} - c^{(i)}_{l}$ and $c^{(i)}_{hh} - c^{(i)}_{h}$ are significantly positive, it would imply that the confidence in guessing others' total effort significantly increases as one moves from Day 1 to Day 2, thus supporting the information motive. We perform four two-sided sign tests to evaluate the hypothesis (two tests per treatment). In all four cases, we fail to reject the null hypothesis of equal confidence (with p-values of 0.79 and 0.69 for Coop-pred and 0.25 and 0.90 for Comp-pred). There is little evidence in favor of the hypothesis that the level of confidence increases when more information becomes available after Day 2. For a majority of forecasters (between 65 and 70 percent), there is no increase at all, leading to the rejection of H3.[29]

To indirectly test H4 and measure strategic sophistication, we additionally asked all predictors to participate in two symmetric two-player games, R and S (see to panel of Figure 3) . Predictors were informed that they would be randomly matched with another predictor, and their joint actions would determine the payment from the game. The two games are identical with just the roles of player 1 and player 2 being switched. In other words, the players basically played in both roles of the same game, in order. In the first game R, there exists a strictly dominant strategy Y for player 1, while in the second game S, player 1 has a dominant strategy B only after she eliminates the strictly dominated strategies X, Z from her partner. Playing correctly in R required rationality (maximizing payoffs based on any belief about the partner), whereas playing correctly in S required strategic reasoning: assuming the partners are rational, and responding to their partner's rationality. Strategically, both games were at least as simple as the game the original workers played under Cooperative or Competitive incentives, thus providing a lower bound on rationality and strategic sophistication. Figure the bottom panel of Figure 3 displays the relative

[28]An additional test fails to reject $e^{(i)}_{l'} = e^{(i)}_{h'}$ when the alternative hypothesis is $e^{(i)}_{l'} > e^{(i)}_{h'}$.

[29]In comparison, predictors react to information about Day 1 versus not knowing anything at all about the workers. We perform four sign tests (two per treatment) in which we test whether $c^{(i)}_{l} - c^{(i)}_{\emptyset}$ and $c^{(i)}_{h} - c^{(i)}_{\emptyset}$ are significantly positive. If they are, that would mean that predictors feel more confident about their guess when some information is available. In all four cases, the tests reject the null hypothesis of equal confidence in favor of the alternatives $c^{(i)}_{l} - c^{(i)}_{\emptyset} > 0$ and $c^{(i)}_{h} - c^{(i)}_{\emptyset} > 0$.

Payoff Matrices

Game R

		Other Mturker		
		X	Y	Z
	A	2.00, 4.50	1.00, 1.00	2.50, 2.00
You	B	4.00, 1.00	2.00, 3.50	3.00, 2.50
	C	3.00, 2.50	1.50, 3.00	2.00, 4.00

Game S

		Other Mturker		
		X	Y	Z
	A	4.50, 2.00	1.00, 4.00	2.50, 3.00
You	B	1.00, 1.00	3.50, 2.00	3.00, 1.50
	C	2.00, 2.50	2.50, 3.00	4.00, 2.00

Joint Distribution of Choices

Coop-pred

		Game S			
		A	B	C	Total
Game R	A	3.06	4.08	3.06	10.20
	B	20.41	14.29	43.88	78.57
	C	5.10	0.00	6.12	11.22
	Total	28.57	18.37	53.06	100.00

Comp-pred

		Game S			
		A	B	C	Total
Game R	A	4.04	5.05	5.05	14.14
	B	23.23	17.17	31.31	71.72
	C	7.07	4.04	3.03	14.14
	Total	34.34	26.26	39.39	100.00

FIGURE 3. Payoff Matrices and Joint Distribution of Choices for Games R and S

frequency of choices in Games R and S for each of the treatments. We observe that while 78% and 72% of predictors choose the optimal strategy for the Coop-pred and Comp-pred treatments respectively, only 18% and 26% of all predictors select the best response to their partner's dominant strategy for those treatments, which is lower than the expected rate under random choice. As a result, it is possible that the few participants who do find the right strategy in S are doing so purely by chance. Among those who do best respond to their partner's dominant strategy, 78% and 65% exhibit rational behavior in the Coop-pred and Comp-pred treatments, respectively.[30] The finding that the conditional rate of getting R right does not improve given that the subject is playing correctly in S suggests that subjects who are getting it right in S are doing so by chance, which supports H4 and indicates that most subjects lack strategic sophistication.

[30]Of the rational predictors, 18 percent and 24 percent best respond to a dominant strategy from the other player in the Coop-pred and Comp-pred treatments, respectively.

Appendix D. Instructions

In this section, we display the instructions for each treatment. The only difference between the instructions for the Cooperative-obs and Competitive-obs treatments with respect to their counterparts in which workers do not observe the daily effort of their partners is that the last sentence, "Neither you, nor your matched partner would ever be able to observe the number of tasks completed by the other person." was replaced by "On each day, both you and your partner would be able to observe the number of tasks completed by the other person, on the previous day(s)."

Instructions

Instructions

In this study, you are asked to log in for **3** consecutive **days** and work on a typing task for up to 45 minutes in total over those 3 dates. For example, if you are reading this on Thursday June 9, then your first login is on June 9, and you are asked to login again on Friday June 10 and Saturday June 11. Each **day**, you may choose how much time you would like to spend completing tasks. You can log in at any time of the day. However, once you login the timer will not stop unless you click on the ***End work for today*** button. Clicking on the red button will allow you to stop and continue on the subsequent day.

Your Job

You will be presented with a series of randomly selected letters. You would be asked to type that exact sequence of letters in the box provided. Note that each character has to be correct. Once you have entered the sequence, click the *Submit* button or press the *Enter* key. Every time you successfully enter two words, you will receive **$0.03/ 2 words (or $0.015 per word)**. You will be able to see the total number of successfully completed tasks by you on the screen.

Sequences must be entered manually (key bindings or automated button-pushing program/scripts cannot be used), or the task will not be approved.

Payment

At the end of the third day, the number of correct words you enter in the 3 tasks will be added up. You will receive a piece-rate of **$0.015 for per word**. In addition to this, if you complete the requirement of logging in each day within the scheduled time and you complete at least the self-selected minimum number of typing tasks, you would guarantee yourself the HIT completion rate of $4.00.

FIGURE 4. *Instructions Solo Treatment*

Instructions

Instructions

In this study, you are asked to log in for **3** consecutive **days** and work on a typing task for up to 45 minutes in total over those 3 dates. For example, if you are reading this on Thursday June 9, then your first login is on June 9, and you are asked to login again on Friday June 10 and Saturday June 11. Each **day**, you may choose how much time you would like to spend completing tasks. You can log in at any time of the day. However, once you login the timer will not stop unless you click on the ***End work for today*** button. Clicking on the red button will allow you to stop and continue on the subsequent day.

Your Job

You will be presented with a series of randomly selected letters. You would be asked to type that exact sequence of letters in the box provided. Note that each character has to be correct. Once you have entered the sequence, click the *Submit* button or press the *Enter* key. Every time you successfully enter two words, you will receive **$0.03/ 2 words (or $0.015 per word)**. You will be able to see the total number of successfully completed tasks by you on the screen.

Sequences must be entered manually (key bindings or automated button-pushing program/scripts cannot be used), or the task will not be approved.

Payment

At the end of the third day, the number of correct words you enter in the 3 tasks will be added up. You will receive a piece-rate of **$0.015 for per word**. In addition to this, if you complete the requirement of logging in each day within the scheduled time and you complete at least the self-selected minimum number of typing tasks, you would guarantee yourself the HIT completion rate of $4.00.

FIGURE 5. *Instructions Cooperative Treatment*

Instructions

Instructions

In this study, you are asked to log in for **3** consecutive **days** and work on a typing task for up to 45 minutes in total over those 3 dates. For example, if you are reading this on Thursday June 9, then your first login is on June 9, and you are asked to login again on Friday June 10 and Saturday June 11. Each **day**, you may choose how much time you would like to spend completing tasks. You can log in at any time of the day. However, once you login the timer will not stop unless you click on the ***End work for today*** button. Clicking on the red button will allow you to stop and continue on the subsequent day.

Your Job and Payment

You will be presented with a series of randomly selected letters. You would be asked to type that exact sequence of letters in the box provided. Note that each character has to be correct. Once you have entered the sequence, click the *Submit* button or press the *Enter* key. Every time you successfully enter two words, you will receive **$0.01 per word**. You will be able to see the total number of successfully completed tasks by you on the screen.
You will be matched with another MTurk participant who is also performing the same tasks. If you are able to perform more tasks than the other player, then you get an additional reward for the total tasks completed by both of you, at the rate of **$0.005 per word**. For example, if the two players completed a total of 200 and 400 words, then the latter player would get an additional reward worth ($0.005 per word x 600 words).
Thus, the additional reward pays the higher-performing player by an amount proportional to the total tasks completed jointly by the two players. If both of you submit the same number of words, you two would share the additional reward equally.
If you log in and complete the self-selected minimum number of tasks in each of the three consecutive days, you would additionally get the HIT completion rate of $4.00.

Sequences must be entered manually (key bindings or automated button-pushing program/scripts cannot be used), or the task will not be approved..

Examples:

At the end of the third day, the number of correct words you enter in the 3 tasks will be added up. The number of correct words the other participant enters during the 3 tasks will also be added up.

Example 1. Suppose that you and the other participant set a minimum of 30 tasks. Suppose also that you enter 105 + 160 + 195 = 460 correct words, and the other participant enters 65 + 205 + 110 = 380 correct words over the three days. In this case, you completed more words than the other participant did. Thus, you will get a payment of

Your Payment

Piece-rate for your own tasks	+	Reward for completing more tasks	Login bonus*	=	Total
460 × $0.01 = $4.60	+	(460 + 380) x $0.005 = $4.20	$4.00	=	$12.80

*In this example the login bonus is earned because the participant completes the minimum number of tasks set by him/her.

The other participant will get:

Other Participant's Payment

Piece-rate for their own tasks	+	Reward for completing more tasks	Login bonus*	=	Total
380 × $0.01 = $3.80	+	$0.00	$4.00	=	$7.80

*In this example the login bonus is earned because the participant completes the minimum number of tasks set by him/her.

Example 2. Suppose that you and the other participant set a minimum of 30 tasks. Suppose also that you enter 55 + 155 + 90 = 300 correct words, and the other participant enters 120 + 120 + 120 = 360 correct words. In this case, you completed fewer words than the other participant did. In this case, your payment and the other participant's payment will be determined as following.

Figure 6. *Instructions Competitive Treatment*

Your Payment

Piece-rate for your own tasks	+	Reward for completing more tasks	Login bonus*	=	Total
300 × $0.01 = $3.00	+	$0.00	$4.00	=	$7.00

*In this example the login bonus is earned because the participant completes the minimum number of tasks set by him/her.

Other Participant's Payment

Piece-rate for their own tasks	+	Reward for completing more tasks	Login bonus*	=	Total
360 × $0.01 = $3.60	+	(300 + 360) x $0.005 = $3.30	$4.00	=	$10.90

*In this example the login bonus is earned because the participant completes the minimum number of tasks set by him/her.

Neither you, nor your matched partner would ever be able to observe the number of tasks completed by the other person.

FIGURE 7. *Instructions Competitive Treatment (Continuation)*

Appendix E. Quiz

uOttawa — Instructions — UCDAVIS UNIVERSITY OF CALIFORNIA

Please answer the following questions. You can always go back and check the instructions using the "Instructions" button above.

You need to answer at least 3 out of the 5 questions correctly to qualify for the HIT.

How do you qualify for the HIT completion and the associated $4 payment?

- By working up to 45 minutes, in total, over the three days (today included)
- By logging into the HIT on each of the three days (today included)
- By logging in to the HIT on each of the three days (today included), AND, completing the minimum number of tasks every day

MTurker Sam plans to spend 20 minutes on the task on the first day, 20 minutes on the second day, and 20 more minutes on the third day. Is his plan realistic?

- Yes
- No, he is planning to spend a total of 60 minutes on the task, and the study allows you only up to 45 minutes on the tasks, in total, over the three days.
- No, he is spending too little time on the tasks

MTurker Jane plans to spend 30 minutes on the task on the first day, 0 minutes on the second day, and 5 more minutes on the third day. Would she qualify for the HIT completion and the associated payment?

- Yes
- No, because she is spending less than 45 minutes in total over the three days
- No, because by not working on the tasks on the second day, she would not be able to complete the minimum number of tasks on that day, which would disqualify her from HIT completion

In the question above about MTurker Jane, what would be her total payment, in case she did not qualify for the HIT completion?

- Zero
- Though she would not get the HIT completion payment, she would still get paid for the typing tasks completed, at the piece-rate, as specified in the instructions.

MTurker Jack spends 45 minutes on the task on the very first day of login, while two more days were left on the study. Which of the following is true?

- Jack would not be able to complete the HIT, as he has exhausted his total 45 minutes, and there is no time left to complete the minimum tasks in days two and three
- Jack's HIT is still on track for completion

Figure 8. *Quiz Solo Treatment*

Appendix F. additional experimental interfaces

In terms of the experimental design, we introduced four features to collect richer data on intertemporal plans and workers' commitment. First, before starting the tasks on Day 1, workers could commit to a minimum number of tasks ($m \geq 5$) that they would have to complete on each of the three days for their participation to count on that day. They could always select $m = 5$ to make it easier to earn the \$4 participation bonus. But by instead setting a higher minimum number, they could also commit to higher effort on future days (Figure 9). Second, workers were asked to submit the number of tasks they planned to complete each day. Adherence to the plan was not incentivized, but the planning feature provided us with non-incentivized data on how workers make and update plans.[31] Third, workers were offered the choice to receive email reminders on Days 2 and 3 to log back into work.[32] Fourth, before starting the typing tasks each day, workers could use an on-screen calculator to calculate their payoff for the number of tasks they and their partner finished. We designed this feature hoping that the calculator entries can potentially serve as non-choice data about subjects' beliefs about the partner. Figure 10 shows the interface of the payoff calculator for a subject in the Cooperative treatment.

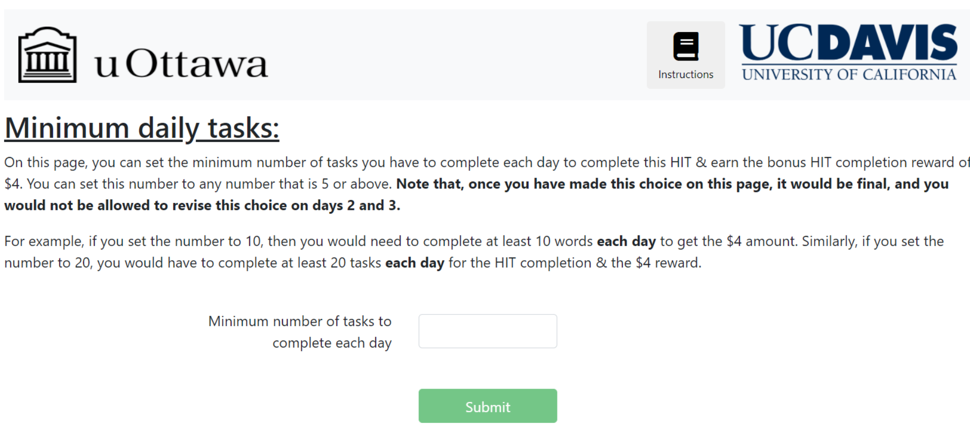


Figure 9. *Commitment Interface*

Figure 11 displays a screenshot illustrating the feedback received by participants regarding their partner's performance within the observability treatments. In this

[31]Unfortunately, workers did not utilize the planning feature as expected and their plans were significantly more conservative than their observed productivity.

[32]About 89 percent of the workers opted-in for the reminders, and thus there wasn't a lot of variation across treatments.

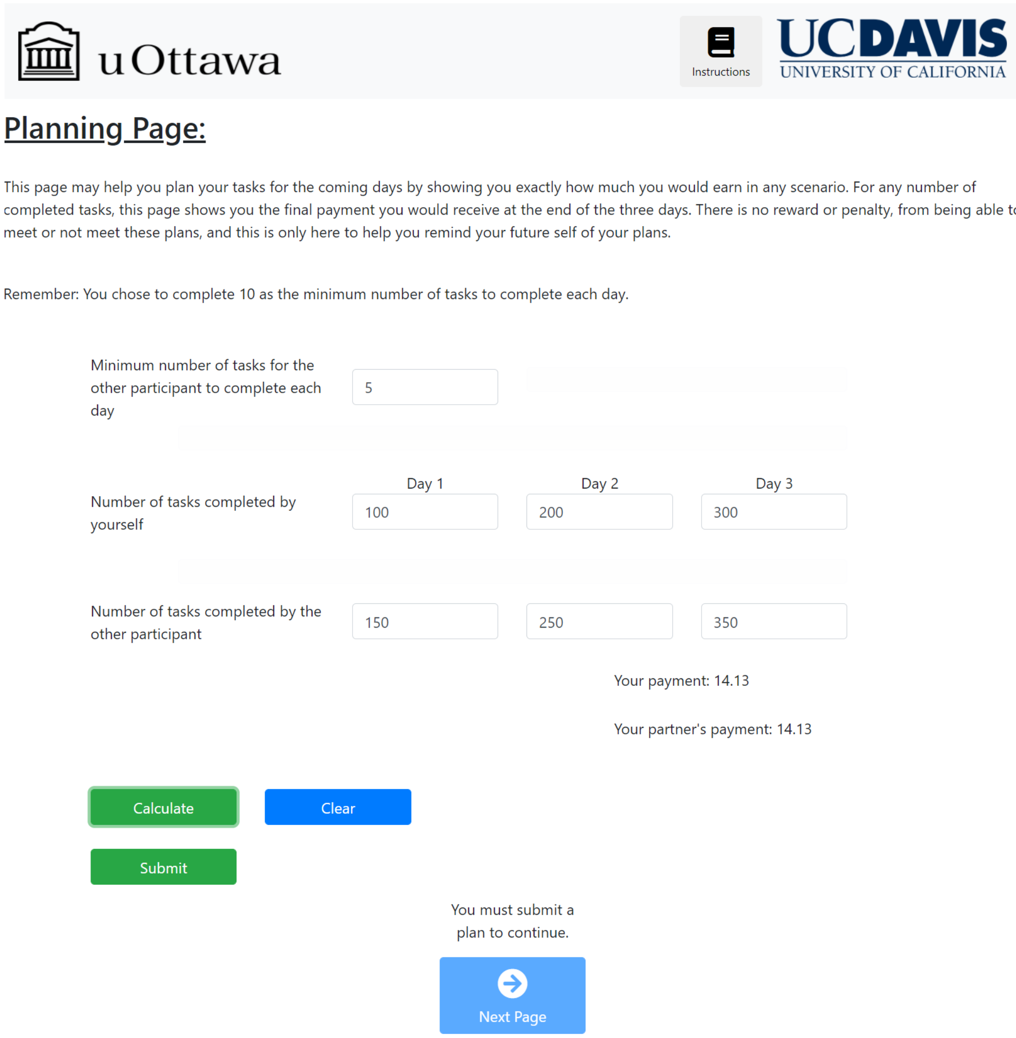


FIGURE 10. *Plan/Calculator Interface*

instance, the participant is situated on Day 3, receiving information about their partner's performance on both Day 1 and Day 2.

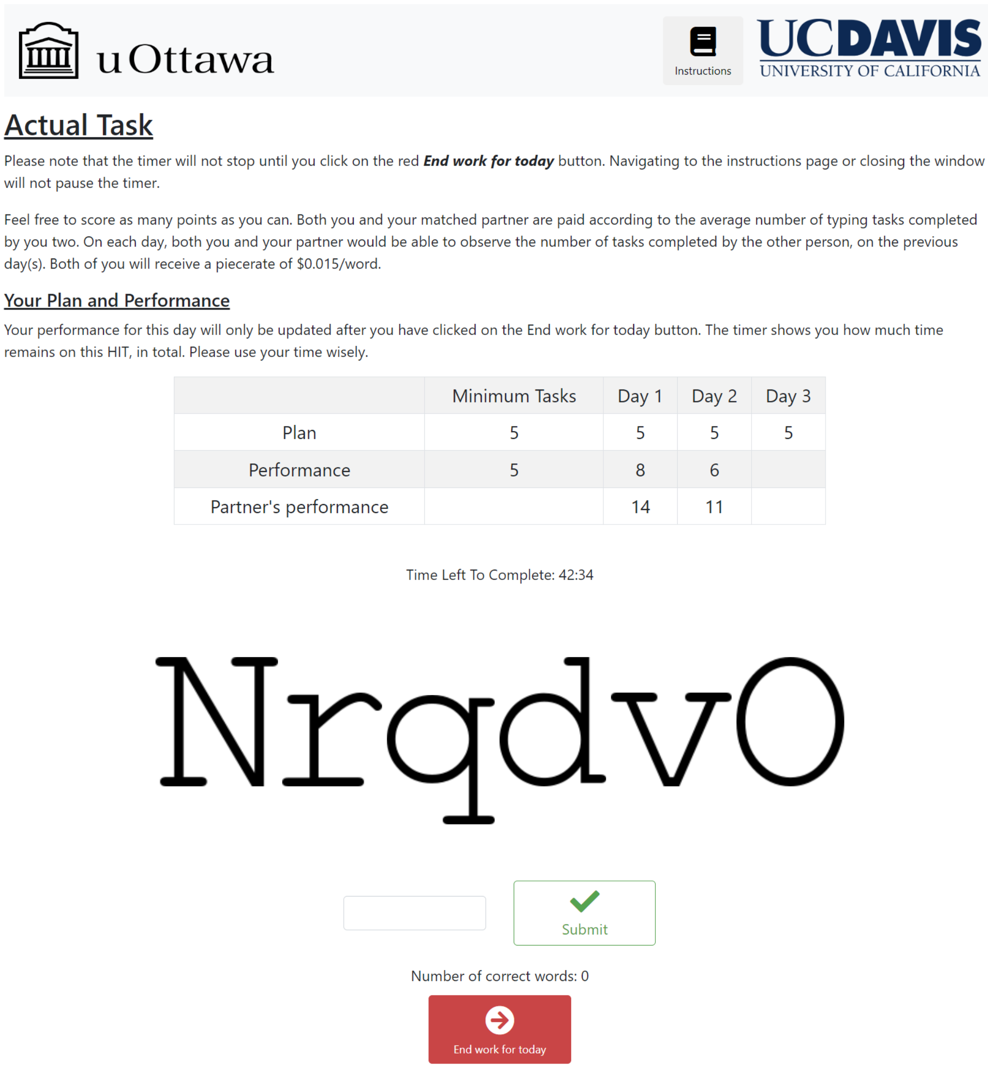


FIGURE 11. *Feedback about the partner's performance*

Appendix G. Instructions for the guessing tasks

Below, we show the instructions for the guessing tasks. Note that the first page corresponds to the Coop-pred treatment, while pages 2 and 3 correspond to the Comp-pred treatment. The rest of the instructions were identical for both treatments.

uOttawa UCDAVIS UNIVERSITY OF CALIFORNIA

Guessing Assignment:

In the past months, a few Mturkers participated in a study that we will refer to as 3-Day Study. Your goal is to make guesses about their performance in 3-Day Study. You can earn up to $4.50 for your guesses based on how accurate they are.

To help you in your predictions, we will show you the following background information:

1. The instructions that were seen by the participants in 3-Day Study (followed by a comprehension quiz)
2. The type of tasks participants in 3-Day Study completed (followed by asking you to complete 10 similar tasks)

Below, we include the instructions from 3-Day Study in italics. Please read the instructions carefully; they will help you guess better and do well on the quiz. Only those who do well on the quiz would qualify to participate in the guessing task.

Instructions seen by past participants from 3-Day Study

In this study, you are asked to log in for 3 consecutive days and work on a typing task for up to 45 minutes in total over those 3 dates.

Your Job and Payment

You will be presented with a series of randomly selected letters. You would be asked to type that exact sequence of letters in the box provided. Note that each character has to be correct. Once you have entered the sequence, click the Submit button or press the Enter key. Each correctly entered sequence counts as completing one task. You will be able to see the total number of successfully completed tasks by you on the screen.

You will be randomly matched with another MTurk participant. At the end of the third day, the number of correct words you enter over the 3 days will be added up. The number of correct words the other participant enters during the 3 days will be also added up. When calculating your payment, we pay both of you for the average number of correct words you and the other participant have entered. Both of you will receive a piece-rate of $0.015 per word.

Let's analyze two examples.

Example 1. *Suppose, you enter 105 + 160 + 195 = 460 correct words, and the other participant enters 65 + 205 + 110 = 380 correct words. In this case, the average is 420. Thus, you and the other participant will get a payment of 420 × $0.015 = $6.30 each.*

Example 2. *Suppose, you enter 55 + 155 + 90 = 300 correct words, and the other participant enters 120 + 120 + 120 = 360 correct words. In this case, the average is 330. Thus, you and the other participant will get a payment of 330 × $0.015 = $4.95 each.*

On each day, both you and your partner would be able to observe the number of tasks completed by the other person, on the previous day(s).

Below, you will find the actual tasks that workers from 3-Day Study had to complete. Please submit 10 tasks to gain experience about their job and proceed to the guessing tasks.

uOttawa

UC DAVIS
UNIVERSITY OF CALIFORNIA

Guessing Assignment:

In the past months, a few Mturkers participated in a study that we will refer to as 3-Day Study. Your goal is to make guesses about their performance in 3-Day Study. You can earn up to $4.50 for your guesses based on how accurate they are.

To help you in your predictions, we will show you the following background information:

1. The instructions that were seen by the participants in 3-Day Study (followed by a comprehension quiz)
2. The type of tasks participants in 3-Day Study completed (followed by asking you to complete 10 similar tasks)

Below, we include the instructions from 3-Day Study in italics. Please read the instructions carefully; they will help you guess better and do well on the quiz. Only those who do well on the quiz would qualify to participate in the guessing task.

Instructions seen by past participants from 3-Day Study

In this study, you are asked to log in for 3 consecutive days and work on a typing task for up to 45 minutes in total over those 3 dates.

Your Job and Payment

You will be presented with a series of randomly selected letters. You would be asked to type that exact sequence of letters in the box provided. Note that each character has to be correct. Once you have entered the sequence, click the Submit button or press the Enter key. Every time you successfully enter two words, you will receive $0.01 per word. You will be able to see the total number of successfully completed tasks by you on the screen. You will be matched with another MTurk participant who is also performing the same tasks. If you are able to perform more tasks than the other player, then you get an additional reward for the total tasks completed by both of you, at the rate of $0.005 per word. For example, if the two players completed a total of 200 and 400 words, then the latter player would get an additional reward worth ($0.005 per word x 600 words).

Thus, the additional reward pays the higher-performing player by an amount proportional to the total tasks completed jointly by the two players. If both of you submit the same number of words, you two would share the additional reward equally.

Let's analyze two examples.

Example 1. *Suppose that you enter 105 + 160 + 195 = 460 correct words, and the other participant enters 65 + 205 + 110 = 380 correct words over the three days. In this case, you completed more words than the other participant did. Thus, you will get a payment of*

Your Payment

Piece-rate for your own tasks	+	***Reward for completing more tasks***	=	***Total***
460 x $0.01	+	*(460 + 380) x $0.005*	=	*$8.80*
$4.60	+	*$4.20*		

The other participant will get:

Other Participant's Payment

Piece-rate for your own tasks	+	***Reward for completing more tasks***	=	***Total***
380 × $0.01	+	*$0.00*	=	*$3.80*
$3.80	+			

Example 2. *Suppose that you enter 55 + 155 + 90 = 300 correct words, and the other participant enters 120 + 120 + 120 = 360 correct words. In this case, you completed fewer words than the other participant did. In this case, your payment and the other participant's payment will be determined as following.*

Your Payment

Piece-rate for your own tasks	**+**	***Reward for completing more tasks***	**=**	***Total***
300 × $0.01	+	*$0.00*	=	*$3.00*
$3.00	+			

The other participant will get:

Other Participant's Payment

Piece-rate for your own tasks	**+**	***Reward for completing more tasks***	**=**	***Total***
360 x $0.01	+	*(300 + 360) x $0.005*	=	*$6.90*
$3.60	+	*3.30*		

On each day, both you and your partner would be able to observe the number of tasks completed by the other person, on the previous day(s).

uOttawa | Instructions | UCDAVIS UNIVERSITY OF CALIFORNIA

Guessing Tasks I

In the table below, you have to guess the performance of Mturkers and their partners on the task described before. These are called guessing tasks.

Remember, on each day, paired MTurkers were able to observe the number of tasks completed by the other person, on the previous day(s).

For each question, if you can guess perfectly, you will get a prize of $0.50. As your guess moves further away from the actual value, this prize will decrease proportionally to the square of the error. Over-guessing or under-guessing reduces your prize in the same proportion. In particular, if the true value is t, and your guess is g, then your error is (g-t). Therefore, you will get $0.50 - (error/100)^2. If your guessing error is more than 70, you will get zero.

Below each guessing task, we also ask you to report your confidence in your guess. Please answer honestly, as your response would help us in the study.

We have randomly chosen **Worker A**, **Worker B**, and **Worker C** from the 3-Day Study. You know nothing about **Worker A**. **Worker B** completed between **20 to 30 tasks on Day 1**. **Worker C** completed between **100 and 110 tasks on Day 1**. Please fill out the following table with your guesses. Below each guess, you must also enter how confident you are in your guess in percentage terms (0 to 100, where 0 means no confidence at all, and 100 means completely confident).

	By **Worker A**	By **Worker B**	By **Worker C**	By **Worker B**'s partner	by **Worker C**'s partner
Guess the Total tasks completed over 3 days					
Your confidence (0-100)					

We have randomly chosen **Worker D** and **Worker E** from the 3-Day Study. **Worker D** completed between **50 to 80 tasks in total over Days 1 and 2.** **Worker E** completed between **220 and 250 tasks in total over Days 1 and 2**. Please fill out the following table with your guesses. Below each guess, you must also enter how confident you are in your guess in percentage terms (0 to 100, where 0 means no confidence at all, and 100 means completely confident).

	By **Worker D**	By **Worker E**
Guess the Total tasks completed over 3 days		
Your confidence (0-100)		

Remember that in 3-Day Study, paired MTurkers were able to observe the number of tasks completed by the other person on the previous day(s). **We have run another study, 3-Day Study-Private, where the rules of payments and working were identical, and the only difference was that paired Mturkers** could not observe **the number of tasks completed by each other.** We will randomly choose Workers G and H from the 3-Day Study and 3-Day Study-Private, respectively. Please guess the number of tasks completed by them **on the first day** of the study.

	By **Worker G (From 3-Day Study)**	By **Worker H (From 3-Day Study-Private)**
Guess the tasks completed only on day 1		
Your confidence (0-100)		

If you had the opportunity to participate in either 3-Day Study or 3-Day Study-Private, which one would you have chosen?

- ○ 3-Day Study
- ○ 3-Day Study-Private

Why?
(Pick any or all that apply)

- ☐ I think I would earn more in the study I chose.
- ☐ I would rather work privately.
- ☐ I would rather not work privately.
- ☐ I would like to know how much fellow co-MTurkers are working.
- ☐ I would like fellow co-MTurkers to know how much I am working.

Decision Tasks For You

- In the following question, you will choose between Options A, B, and C.
- We will match you with another MTurker who is completing an identical survey.
- The other Mturker will choose between X, Y and Z.
- Both of you will decide independently, without knowing each other's choice.

Your and the other Mturker's payments will depend on the combination of your and their simultaneous choices. These payment possibilities are represented in the table below

		Other Mturker		
		X	Y	Z
You	A	**$2.00,** $4.50	**$1.00,** $1.00	**$2.50,** $2.00
	B	**$4.00,** $1.00	**$2.00,** $3.50	**$3.00,** $2.50
	C	**$3.00,** $2.50	**$1.50,** $3.00	**$2.00,** $4.00

Notice that the cells contain a pair of dollar earnings. In each cell, the first dollar earning (in bold) shows your payment, and the second number indicates the payment of the other Mturker.

Your choice (along the row) and the other worker's choice (along the column) jointly determine the cell that is relevant for payment. For example, if you choose C, and the other Mturker chooses X, then you get $3.00 and the other Mturker gets $2.50.

Please respond on the next quiz to proceed to the next page:

Quiz

1. How much would you earn if you select B, and the other Mturker chooses X?

2. How much would you earn if you select C, and the other Mturker chooses Y?

3. How much would the other Mturker earn if you select C, and the other Mturker chooses Z?

4. How much would the other Mturker earn if you select A, and the other Mturker chooses Y?

The table below shows the payoff possibilities for your and the other Mturker's actions. Please choose an action:

		Other Mturker		
		X	Y	Z
You	A	**$2.00,** $4.50	**$1.00,** $1.00	**$2.50,** $2.00
	B	**$4.00,** $1.00	**$2.00,** $3.50	**$3.00,** $2.50
	C	**$3.00,** $2.50	**$1.50,** $3.00	**$2.00,** $4.00

- ○ A
- ○ B
- ○ C

uOttawa

Instructions

UCDAVIS UNIVERSITY OF CALIFORNIA

Question 2

In this task, you will choose between Options A, B, and C as before, but the payoff possibilities are different. The other Mturker will also choose between X, Y and Z independently.

		Other Mturker		
		X	Y	Z
You	A	**$4.50,** $2.00	**$1.00,** $4.00	**$2.50,** $3.00
	B	**$1.00,** $1.00	**$3.50,** $2.00	**$3.00,** $1.50
	C	**$2.00,** $2.50	**$2.50,** $3.00	**$4.00,** $2.00

In each cell, the first payoff (in bold) shows your payment, and the second number indicates the payment of the other Mturker. Your choice and the other worker's choice determine the cell is relevant for payment.

Your Choice

- ○ A
- ○ B
- ○ C

Appendix H. Theory Results and Proofs

Proof of Proposition 1:

Proof. Part (i): As defined in the proposition, $c^\gamma = \beta\delta\frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}} + \beta\delta^2\frac{1}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}$.
The candidate solutions for e_1 are

$$e_1^{s*} \quad = \quad \begin{cases} (\beta\delta^t pk)^{\frac{1}{\gamma}} & \text{if } \beta \leq \beta_1, M = (\beta_1\delta^t kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-1}kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-2}kp)^{\frac{1}{\gamma}} \\ M - (\hat{\beta}\delta^{t-2}pk)^{\frac{1}{\gamma}} - (\hat{\beta}\delta^{t-1}pk)^{\frac{1}{\gamma}} & \text{if } \beta \in [\beta_1, \beta_2], M = (1+(\beta_2\delta)^{1/\gamma})(\hat{\beta}\delta^{t-1}kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-2}kp)^{\frac{1}{\gamma}} \\ \frac{(\beta\delta)^{\frac{1}{\gamma}}(M-(\hat{\beta}\delta^{t-2}pk)^{\frac{1}{\gamma}})}{1+(\beta\delta)^{\frac{1}{\gamma}}} & \text{if } \beta \in [\beta_2, \beta_3] \text{ where } \beta_3 \text{ equates this expression with } \frac{c}{1+c}\text{M} \\ \frac{c}{1+c}M & \text{if } \beta \geq \beta_3 \end{cases}$$

We omit the more technical details of the derivation for brevity. Fix some M, δ, k, p: The candidate solution depends on how high the present bias β is, and can take four values based on which range β lies in. The ranges are bounded by the values $\beta_1, \beta_2, \beta_3$ which are themselves implicitly defined next to the ranges. For example, in the lowest possible range of $\beta \leq \beta_1$, the worker should complete $(\beta\delta^t pk)^{\frac{1}{\gamma}}$ tasks. And β_1 is implicitly defined by the equation

$$M = (\beta_1\delta^t kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-1}kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-2}kp)^{\frac{1}{\gamma}}$$

At $\beta = \beta_1$, we are at the intersection of two ranges, and their corresponding optimal solutions are exactly equal *by definition of* β_1. To see this,

$$\begin{aligned} M &= (\beta_1\delta^t kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-1}kp)^{\frac{1}{\gamma}} + (\hat{\beta}\delta^{t-2}kp)^{\frac{1}{\gamma}} \\ \iff M - (\hat{\beta}\delta^{t-2}pk)^{\frac{1}{\gamma}} - (\hat{\beta}\delta^{t-1}pk)^{\frac{1}{\gamma}} &= (\beta_1\delta^t pk)^{\frac{1}{\gamma}} \end{aligned}$$

The continuity of optimal solutions holds between every two ranges.
Ceteris paribus, as β increases, there are two effects. First, each expression of maximal e_1^{s*} weakly increases for any fixed value of p that lies in the interior of any of the four ranges $\beta \leq \beta_1$, $\beta \in [\beta_1, \beta_2]$, $\beta \in [\beta_2, \beta_3]$, $\beta \geq \beta_3$. Thus, while comparing within the same range, e_1^{s*} must increase with β. Second, the ranges themselves depend on β. To account for this, take $\beta_l < \beta_h$ such that they lie on two sides of a boundary, say β_1: thus $\beta_l < \beta_1 < \beta_h < \beta_2$. Then, using the conotinuity of candidate solutions at β_1, we get $e_1^{s*}|_{\beta=\beta_l} < e_1^{s*}|_{\beta=\beta_1} = e_1^{s*}|_{\beta=\beta_h}$. Thus, the monotonicity holds regardless.

The (ii) On Day 3, the individual is willing to work up to e_3^{max} tasks where the marginal cost catches up to the marginal benefit

$$\frac{d}{de_3}\beta\delta^{t-2}kp\,(e_1 + e_2 + e_3) = \frac{d}{de_3}\frac{(e_3)^{\gamma+1}}{\gamma+1}$$

By experimental rules, she is allowed to work up to $M - e_1 - e_2$ tasks. Thus she stops at the number that is smaller among these two.

Step 1: Setting up the Lagrangean for Day 2,

$$\mathcal{L}_2 \;=\; \beta\delta^{t-1}kp\,(M) - \frac{(e_2)^{\gamma+1}}{\gamma+1} - \beta\delta\frac{(M-e_1-e_2)^{\gamma+1}}{\gamma+1})$$

Because she believes that she would complete all M tasks, we use M as the total number of tasks completed. The necessary first order conditions are

$$-(e_2)^{\gamma} + \beta\delta(M-e_1-e_2)^{\gamma} \;=\; 0$$

We get $-e_2^{\gamma} + \beta\delta(M-e_1-e_2)^{\gamma} = 0$, i.e, $e_2^{s*} = \frac{(\beta\delta)^{\frac{1}{\gamma}}}{1+(\beta\delta)^{\frac{1}{\gamma}}}(M-e_1)$. And she plans to complete $M - e_1 - e_2$ tasks on day 3.

Step 2: Setting up the Lagrangean for Day 1, and substituting the values of e_2^s and e_3^s

$$\mathcal{L}_1^3 \;=\; \beta\delta^t kp(M) - \frac{e_1^{\gamma+1}}{\gamma+1} - \beta\delta\frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}(M-e_1)^{\gamma+1}}{(1+(\hat{\beta}\delta)^{\frac{1}{\gamma}})^{\gamma+1}(\gamma+1)} - \beta\delta^2\frac{(M-e_1)^{\gamma+1}}{(1+(\hat{\beta}\delta)^{\frac{1}{\gamma}})^{\gamma+1}(\gamma+1)}$$

The necessary conditions for maximization are:

$$\beta\delta\frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}(M-e_1)^{\gamma} + \beta\delta^2\frac{1}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}(M-e_1)^{\gamma} \;=\; e_1^{\gamma}$$

Let $c^{\gamma} = \beta\delta\frac{(\hat{\beta}\delta)^{\frac{\gamma+1}{\gamma}}}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}} + \beta\delta^2\frac{1}{((\hat{\beta}\delta)^{\frac{1}{\gamma}}+1)^{\gamma+1}}$, and then, $e_1^{s*} = \frac{c}{1+c}M$. $\square$

Proposition 2. *Consider the Cooperative treatment with piece rate p. Under selfish preferences, a subject has a dominant strategy to exert an effort equal to the Solo-optimal at piece rate $p/2$, i.e, $e_t^{coop*}(p) = e_t^{s*}(p/2)\forall t$. Under an intrinsic motivation parameter $\alpha \in (0,1)$, she has a dominant strategy to exert an effort equal to the Solo-optimal at piece rate $p(1+\alpha)/2 \in [p/2, p)$, i.e, $e_t^{coop*}(p) = e_t^{s*}\,(p(1+\alpha)/2)\,\forall t$.*

Proof. i takes into account both her own and partner's utility. For i, Day 3's problem becomes:

$$e_3^{coop}(e_1,e_2,\beta,p) \equiv \arg\max_{e_3} k\beta\delta^{t-2}\left[\underbrace{\frac{p}{2}(\sum_{t=1}^{3} e_t^{coop} + \boldsymbol{E_b}\sum_{\boldsymbol{t=1}}^{\boldsymbol{3}} \boldsymbol{e'_t})}_{\text{own pay}} + \alpha\underbrace{\frac{p}{2}(w\sum_{t=1}^{3} e_t^{coop} + \boldsymbol{E_b}\sum_{\boldsymbol{t=1}}^{\boldsymbol{3}} \boldsymbol{e'_t})}_{\text{partner's pay}}\right]$$

$$- \left(c(e_3^{coop}) + \alpha \underbrace{\boldsymbol{E_b c(e_3)}}_{\text{partner's cost}, \frac{d(\cdot)}{de_3}=0}\right)$$

$$\text{subject to} \quad \underbrace{0 \le e_3^{coop} \le M - e_1^{coop} - e_2^{coop}}_{\text{identical to Day 3 constraint from Solo}}$$

For simplicity we have replaced the cost of effort with with the function $c(e) = \frac{e^{\gamma+1}}{\gamma+1}$. The constraint remains identical to that in the Solo treatment, and after differentiation, the Langangian takes the same form as in the Solo treatment, with a piece rate of $p(1+\alpha)/2$, representing the marginal return per task. Consequently, the optimal response for i is to behave as if they are in the Solo treatment, receiving a piece rate of $p(1+\alpha)/2$. The partner's effort and effort-cost variables no longer affect the derivatives. Therefore, the best response for player i, denoted as $e_3^{coop}(e_1,e_2,p) = e_3^s(e_1,e_2,\beta,p(1+\alpha_1)/2)$, becomes *independent* of the partner's choice, making it a dominant strategy in this interaction. By solving backwards,
Day 2's problem:

$$e_2^{coop}(e_1,\beta,p) \equiv \arg\max_{e_2} k\beta\delta^{t-1}\left[\frac{p}{2}(1+\alpha)\left(\sum_{t=1}^{3} e_t^{coop} + \boldsymbol{E_b}\sum_{\boldsymbol{t=1}}^{\boldsymbol{3}} \boldsymbol{e'_t}\right)\right]$$

$$- \left(c(e_2^{coop}) + \alpha \underbrace{\boldsymbol{E_b c(e'_2)}}_{\text{partner's cost}, \frac{d(\cdot)}{de_3}=0}\right)$$

$$- \left(\beta\delta c(\hat{e}_3^{coop}) + \alpha\beta'\delta' \underbrace{\boldsymbol{E_b c(e'_3)}}_{\text{partner's cost}, \frac{d(\cdot)}{de_3}=0}\right)$$

subject to

$$\begin{aligned} &(i) \quad 0 \leq e_2^{coop} \leq M - e_1^{coop} \\ &(ii) \quad \underbrace{\hat{e}_3^{coop} \equiv e_3^s(e_1, e_2, \hat{\beta}; p(1+\alpha)/2)}_{\text{identical to Day 2 constraint from Solo}} \end{aligned}$$

The bolded terms representing the partner's effort disappear from the first-order condition, resulting in a Lagrangian that is identical to the one obtained from the Solo treatment, with a piece rate of $p(1+\alpha)/2$. Therefore, by definition, $e_2^{coop}(e_1, \beta, p) = e_2^s(e_1, \beta, p(1+\alpha)/2).)$.

When solving backwards for Day 1, the same steps demonstrate that the marginal benefit per task remains $p(1+\alpha)/2$. Additionally, the constraints, as well as the partner's effort or effort cost, become irrelevant to the optimal response, thereby establishing a dominant strategy. $\square$

Appendix I. Additional Regressions

The following regressions are equivalent to those shown in Table 4, but with the addition of demographic controls.

Table 10. Effect of Partner's past Effort with Demographic Controls

	Dependent variable: Number of Tasks on Day 2			
	Cooperative	Competitive	Cooperative-obs	Competitive-obs
	[1]	[2]	[3]	[4]
$e_{j,1}$	0.096	-0.182^{*}	0.102	0.361^{**}
	(0.09)	(0.10)	(0.07)	(0.15)
$e_{i,1}$	0.222^{*}	0.410^{**}	0.531^{***}	0.144
	(0.11)	(0.18)	(0.10)	(0.17)
$(1_{e_{i,1}>e_{j,1}}) \times e_{j,1}$	0.042	-0.069	0.198	0.453^{**}
	(0.13)	(0.25)	(0.13)	(0.22)
Constant	65.378^{**}	121.595^{***}	48.031	70.294^{**}
	(29.80)	(28.29)	(31.79)	(32.17)
Controls	Yes	Yes	Yes	Yes
Observations	94	96	94	92

Note: Standard errors, reported in parentheses, are robust across all treatments. For treatments with observability, they are clustered at the team level. In our notation, $e_{j,1}$ denotes the number of tasks submitted by the partner on Day 1, while $e_{i,1}$ represents the worker's own effort on the same day. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. Workers that could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The following regressions are equivalent to those shown in Table 5, but with the addition of demographic controls.

Table 11. How Observability Influences Team Treatments?

	Dependent variable			
	Total No. of Tasks	Tasks on D1	Tasks on D2 + D3	1=Spent 45 mins 0=Otherwise
	[1]	[2]	[3]	[4]
Cooperative	445.827***	147.784***	306.464***	0.686***
	(41.94)	(17.21)	(32.91)	(0.12)
Competitive	414.623***	146.788***	273.756***	0.593***
	(47.21)	(17.20)	(36.58)	(0.12)
Cooperative×obs	-24.058	-2.446	-23.177	-0.042
	(24.18)	(9.82)	(19.04)	(0.07)
Competitive×obs	39.767	-1.378	38.953*	0.090
	(26.84)	(9.59)	(21.97)	(0.07)
Controls	Yes	Yes	Yes	Yes
Observations	414	442	414	414

Note: Robust standard errors are reported in parentheses. All specifications are OLS regressions. In every column, except for [2], workers from the Cooperative and Competitive treatments who submitted tasks on Day 1 but could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

The following regressions compare the performance in the Solo treatment relative to Cooperative-obs and Competitive-obs.

Table 12. Cooperative-obs vs Competitive-obs

	Dependent variable			
	Total No. of Tasks	Tasks on D1	Tasks on D2 + D3	1=Spent 45 mins 0=Otherwise
	[1]	[2]	[3]	[4]
Cooperative obs	44.286*	23.715***	22.721	0.136**
	(25.77)	(8.36)	(20.71)	(0.07)
Competitive obs	81.880***	23.317***	56.975***	0.194***
	(25.83)	(8.48)	(21.70)	(0.06)
Constant (Solo)	344.663***	121.703***	229.171***	0.444***
	(42.05)	(13.92)	(33.09)	(0.11)
Controls	Yes	Yes	Yes	Yes
Observations	330	348	330	330
	p- values from F tests			
Cooperative obs vs Competitive obs	0.16	0.96	0.12	0.37

Note: Robust standard errors are reported in parentheses. All specifications are OLS regressions. In every column, except for [2], workers from the Cooperative and Competitive treatments who submitted tasks on Day 1 but could not be paired with another worker due to an odd number of subjects in the session (and consequently could not participate on subsequent days) are also excluded. It should be noted that workers who exhausted the full 45 minutes on either Day 1 or Day 2 are excluded from this analysis. The controls include dummy variables on gender, age (dummy variables for ages 18 to 24, 25 to 30, 31 to 40, 41-50, 51-64 and >65), and the education level (dummy variables for subjects that completed high school or less, some college or a bachelor's degree or more). $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.