

Cerrone, Claudia; Chakraborty, Anujit; Kim, Hyok Jung; Lades, Leonhard K.

**Working Paper**

## Estimating Present Bias and Sophistication over Effort and Money

Working Paper, No. 359

**Provided in Cooperation with:**

University of California Davis, Department of Economics

*Suggested Citation:* Cerrone, Claudia; Chakraborty, Anujit; Kim, Hyok Jung; Lades, Leonhard K. (2023) : Estimating Present Bias and Sophistication over Effort and Money, Working Paper, No. 359, University of California, Department of Economics, Davis, CA

This Version is available at:

<https://hdl.handle.net/10419/283203>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Estimating Present Bias and Sophistication over Effort and Money <sup>\*</sup>

Claudia Cerrone<sup>†</sup> Anujit Chakraborty<sup>‡</sup> Hyok Jung Kim<sup>§</sup> Leonhard Lades<sup>¶</sup>

November 12, 2023

## Abstract

We design and implement the first real-effort experiment that can jointly estimate present bias ( $\beta$ ) and sophistication ( $\hat{\beta}$ ), with *separate preference parameters* for money ( $\beta_m, \hat{\beta}_m$ ) and effort ( $\beta_e, \hat{\beta}_e$ ). In our study, participants chose to (and predicted to) complete 14% (and 10%) *fewer* tasks on the same day than on a future day, leading to an estimated  $\beta_e$  between 0.70 and .79 (and  $\hat{\beta}_e$  between 0.80 and .88). Participants chose to (and predicted to) complete 2% (and 2%) fewer tasks when the payment happened on the same day than on a future date, implying insignificant present bias or sophistication over money. At the individual-level, estimates of  $\beta_e$  and  $\hat{\beta}_e$  are highly correlated but that of  $\beta_m$  and  $\hat{\beta}_m$  are uncorrelated. We explain and illustrate how assuming identical preferences for money and effort, biases the estimates of key preference parameters.

**JEL Codes:** C91, D81, D90.

**Keywords:** time preferences; present bias; sophistication.

---

<sup>\*</sup>We thank David Freeman and the audience at the [VEABES](#) for comments and feedback.

<sup>†</sup>City, University of London, Email: [claudia.cerrone@city.ac.uk](mailto:claudia.cerrone@city.ac.uk)

<sup>‡</sup>University of California, Davis, Email: [chakraborty@ucdavis.edu](mailto:chakraborty@ucdavis.edu)

<sup>§</sup>Korea Institute for International Economic Policy, Email: [hjkim@kiep.go.kr](mailto:hjkim@kiep.go.kr)

<sup>¶</sup>University College Dublin, Email: [leonhard.lades@ucd.ie](mailto:leonhard.lades@ucd.ie)

# 1 Introduction

*Present bias*, the desire for immediate gratification, often jeopardizes our long-term goals. Individuals who anticipate their present-bias, sometimes plan around it by self-imposing stringent deadlines for their own work (Ariely and Wertenbroch, 2002), investing in illiquid savings products (Ashraf et al., 2006), choosing dominated contracts that offer a bonus wage only when they achieve a high productivity threshold (Kaur et al., 2015), or by investing in commitment devices that help abstain from harmful habits (Schilbach, 2019).<sup>1</sup> Thus, the consumption decisions made by present biased individuals depend on the extent of their present bias, as well as their ability to anticipate their own present bias (Phelps and Pollak, 1968).

This paper proposes and implements a novel experimental design that can jointly estimate discounting, present bias, and the ability to anticipate present bias at the individual level. Following O'Donoghue and Rabin (1999, 2001), we call the ability to anticipate present bias as *sophistication*. In our three-day experiment, participants choose the number of time-dated real-effort tasks that they want to complete for different wages on the current day and on future days. Moreover, we also ask participants to predict the number of tasks that their future selves will choose to complete in the future, and we use these predictions to estimate their sophistication.

Three design features separate this study from other time-preference experiments. First, while most of the existing elicitation literature focuses on estimating only present bias using Convex Time Budgets (Andreoni and Sprenger, 2012; Augenblick et al., 2015; Andersen et al., 2008; Halevy, 2015) or “money earlier versus money later” binary choices (Andersen et al., 2008; Halevy, 2015), we use an effort-choice and prediction paradigm (Augenblick and Rabin, 2019) to estimate both present bias and sophistication.

Second, contrary to Augenblick and Rabin (2019), we need not presume that participants have identical present bias over effort and money. We vary if the time-dated work had to be completed on the same day or not, *and* if it was paid for on

---

<sup>1</sup>Relatedly, Toussaert (2018) shows that experimental participants who anticipate self-control costs as in Gul and Pesendorfer (2001), may commit to a smaller choice set even when they expect to resist temptation.

the same day or not. This double-variation (which separates our design from [Augenblick and Rabin, 2019](#)) allows us to estimate the present bias and sophistication over real-effort and money domains *separately*. Thus, if present bias is disproportionately higher for time-dated effort than for time-dated money as hypothesized by a recent theoretical and experimental literature (see Section 2.1), then our identification strategy would not conflate them together and instead lead to more precise estimates of preferences. In Section 3.1, we explain and hypothesize how a misspecified model that wrongly assumes identical present bias under effort and money would bias the preference parameters estimated in studies like ours. Our data confirms our hypotheses.

Finally, we introduce a novel incentive structure under which the marginal cost of effort increases quicker than in other comparable real-effort experiments. Conceptually, this guarantees that the net utility is more concave in effort and thus further aligns the choice scenario with the canonical identification assumption that the observed choices are the *interior maxima* of a participant’s *concave* utility maximization problem. Empirically, this also makes it less appealing to choose the highest amount of effort possible and thus decreases the number of corner solutions chosen by participants. For example, our data contains only 38% corner choices, as compared to 70% corner choices observed under the popular Convex Time Budget method implemented in [Andreoni and Sprenger \(2012\)](#).

We present non-parametric results as well as results from a structural analysis in Section 4. Our non-parametric analysis provides significant evidence for present bias over the effort domain, but much weaker evidence for present bias over the money domain: Controlling for the distance between the payment and effort dates, participants allocate on average 4.17 (14%) more real-effort tasks to future workdays than to the present, and allocate 0.5 (2%) fewer real-effort tasks when the task-payment occurs in the future rather than in the present. Only the former effect is statistically significant. When predicting their own future decisions, participants believe that they will allocate 3.09 (instead of 4.17) more real effort tasks to future workdays, showing statistically significant evidence of partial sophistication over effort.

For our structural analysis, we use the quasi-hyperbolic discounting model

(Phelps and Pollak, 1968; Laibson, 1997) due to its simplicity and tractability.<sup>2</sup> Under this model, the relative discounting between the present utility and the next day's utility is  $\beta\delta$ , where  $\delta$  is the discounting between any two consecutive future days, and  $\beta \leq 1$  is a present bias parameter that ensures that the relative discounting between two consecutive periods is sharpest between the present and the next day. The smaller is  $\beta$ , the higher the discrepancy between immediate and long-run discounting, and the higher the extent of present bias. Sophistication is measured as the point belief  $\hat{\beta}$  that individuals hold about their future self's present bias parameter. Across three different specifications, on the aggregate level, we structurally estimate a present bias for effort between .70 and .79, and a present bias for money between .95 and 1.09. Irrespective of the specification, the present bias over effort is consistently and significantly different from 1, whereas the present bias over money is never significantly different from 1.

Consistent with previous studies (Andreoni and Sprenger, 2012; Andreoni et al., 2015; Augenblick et al., 2015), we observe little present bias in the money domain but significantly more present bias in the effort domain. These results validate our identification strategy of allowing different present biases over the effort and money domains.

The aggregate-level structural estimates also show that participants are at least partially aware of their present bias in the effort domain. The point estimate of  $\hat{\beta}_e$  is between 0.80 and .88 across specifications and lies between  $\beta_e$  and 1 for the corresponding specification.

We re-estimate our data under the restrictive assumption of identical money and effort preferences. This counterfactual exercise confirms our hypotheses (Section 3.1) about how this specification biases the estimates of preference parameters. For example, we find that MLE tries to fit this misspecified model to the data by estimating a  $\beta$  and  $\hat{\beta}$  that are mutually closer to each other (and smaller in magnitude) than under our other specifications, thus overestimating relative sophistication.

The individual-level structural estimates show that over 67 percent of the par-

---

<sup>2</sup>See Chakraborty (2021) for a detailed discussion of the many ways present-biased preferences can be modeled.

ticipants are present-biased over effort, while only 48 percent of the participants are present-biased over money. The structural estimates confirm that the present bias over effort and present bias over money are strongly and positively correlated, as also observed in our non-parametric analysis. This is also consistent with [Cheung et al. \(2022\)](#), who find a correlation between present bias over money and present bias over food.

The correlation between the individual-level estimates of  $\beta_e$  and  $\hat{\beta}_e$  are also highly correlated (coefficient = 0.96,  $p < .001$ ) and much higher than the range 0.24-0.28 reported in [Augenblick and Rabin \(2019\)](#). The lack of significant correlation between  $\beta$  and  $\hat{\beta}$  in [Augenblick and Rabin \(2019\)](#) could be partly explained by the lack of significant correlation between  $\beta_m$  and  $\hat{\beta}_m$  (coefficient = -0.06,  $p = .58$ ) that would have been conflated in [Augenblick and Rabin \(2019\)](#)'s analysis.

We run our design over three consecutive days as a proof of concept of our identification method, even though our study can be easily extended to multiple days or weeks. Three-day studies like ours are less expensive to run, less vulnerable to attrition, and can be easily replicated by applied researchers who intend to elicit present bias and sophistication as explanatory variables that can further explain choices related to savings behavior, illiquid investments, purchase of gym memberships or health insurance, goal-setting at work, personality traits, later life outcomes, etc.

The remainder of the paper is structured as follows. Section 2 describes the experimental design. Section 3 introduces the model and explains the identification strategy. Section 4 presents our results. Section 5 concludes.

## 2 Experimental Design

### 2.1 Overview: Identifying Present bias from choices

In any economic model, discounting is generally applied to dated utility defined over dated consumption. Thus, the trade-off between dated rewards observed by an analyst can reveal present bias only if those dated rewards correspond appropriately with dated utilities. Real-effort experiments including the current study use two (mutually linked) dated rewards: effort and money.

There are two reasons why dated effort (compared to dated money) might link differently with dated utility: First, while experimenters can enforce that the time-dated effort is fully exerted (consumed) at the pre-determined date, they cannot enforce that the time-dated monetary payment is fully consumed on the pre-determined date. Thus, dated effort shares a closer link with the consumption and utility on that date.<sup>3</sup> Second, as [Coller and Williams \(1999\)](#); [Cubitt and Read \(2007\)](#) argue, choices over time-dated payments are more likely to be influenced by factors extraneous to time preferences, for example, by the arbitrage opportunities offered by the interest rates that exist outside of the experimental environment.<sup>4</sup> Overall, unless one assumes a *consume-on-receipt* model where monetary payments produce immediate one-for-one consumption at the date of receipt, one should expect that preferences over time-dated money would be different from preferences over time-dated effort ([Cohen et al., 2020](#)).<sup>5</sup> In this case, presuming their equality would conflate their effects and lead to a misspecified identification exercise.

Instead, we jointly estimate these parameters *separately* for effort and money. To achieve this we ask participants to work over three days, and experimentally vary the time to effort and the time to payment (see section 2.2.2 for details). For a fixed delay to payment, we compare same-day and future-day work allocations to identify present bias over effort. For a fixed delay to effort, we compare work allocations under same-day payment versus future-day payment to identify present bias over money. The study that is most closely related to this paper is [Augenblick and Rabin \(2019\)](#), which also elicits present bias and sophistication, but does not have this dual variation in effort/ payment delay and instead assumes identical

---

<sup>3</sup>A related criticism applies to both time-dated money and effort, and almost all experimental studies of intertemporal choice. Under a life-cycle model, a particular payment/effort should change the marginal consumption/effort on all dates and not just on the date of implementation. This means that a reward can also influence the utility experienced before or after the reward, thus weakening the link between dated reward and dated utility.

<sup>4</sup>For a counter-argument, see [Andreoni et al. \(2018\)](#) who test for arbitrage over dated money payments directly in their experiment and find no evidence for it, thus supporting money as a valid reward medium to measure discounting.

<sup>5</sup>Consistent with this interpretation, [Augenblick et al. \(2015\)](#) find evidence of present bias only in the effort domain, but not in the money domain, when they measure time preferences for the same individual in both domains.

discounting over effort and money.<sup>6</sup>

Finally, to identify preference parameters from the experimental data, we use a structural estimation. For such an exercise, one has to interpret the observed choices being generated from the interior optima of an individual's utility maximization exercise. To minimize the frequency of corner observations that can violate this identification condition, we introduce a novel incentive scheme (see subsection 2.2.1) that disincentivizes corner choices by ensuring an increasing marginal cost of effort.

## 2.2 Experimental Task and Decisions

Each session lasted three days over which participants chose (and predicted) the number of time-dated real-effort tasks that they wanted to complete for different wages on the current day and on future days. Over the three study days, participants saw 9 decision screens similar to the one shown in Figure 1.

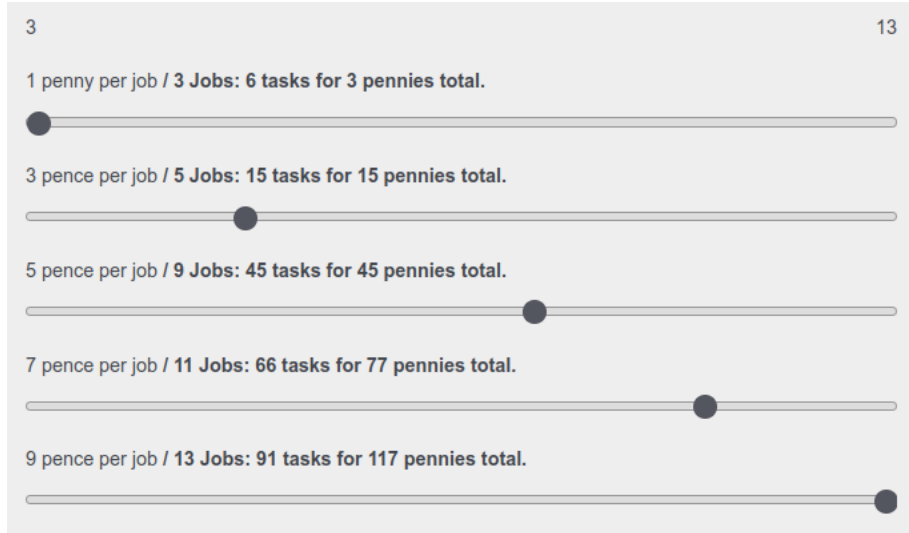


Figure 1: Decision screen for choosing jobs (and tasks) at 5 wages. The screen also specified when the task had to be completed.

On each of the 9 screens, participants saw 5 different wages or piece rates (1p,

---

<sup>6</sup>See equation (1) in [Augenblick and Rabin \(2019\)](#) where their utility model assumes (and estimates) a common  $\beta$  for money or effort. Compare this to our model in Section 3 that assumes different  $\beta$  for money or effort. To the best of our knowledge, [Augenblick and Rabin \(2019\)](#) is the only other paper that directly elicits sophistication as defined in [O'Donoghue and Rabin \(1999\)](#).



3p, 5p, 7p, and 9p) and had to choose the number of time-dated “jobs” (between 3 and 13) that they wanted to complete (on 6 choice screens) or predicted that their future selves would choose to complete (on 3 prediction screens) at that wage and on that date.<sup>7</sup> Jobs translated into real-effort tasks at non-linear rate:  $J$  jobs required completing  $J(J + 1)/2$  real-effort transcription tasks (just *tasks*, henceforth). For example,  $J = 3$  jobs required completing  $3(3 + 1)/2 = 6$  tasks,  $J = 5$  jobs required completing  $5(5 + 1)/2 = 15$  tasks, and  $J = 13$  jobs required completing  $13(13 + 1)/2 = 91$  tasks.

As illustrated by Figure 2, a (transcription) task consisted of typing a sequence of characters into an empty box. If the characters were not typed correctly, an error message appeared and the participant had to retype the characters again.



Figure 2: Example of a transcription task. Completing  $J$  jobs meant finishing  $J(J + 1)/2$  such transcription tasks for a total payment of  $Jp$

### 2.2.1 Increasing marginal cost of jobs

Participants were paid the wage or piece rate *based on the number of jobs completed*. Thus, though every job offered the same marginal benefit of  $p$ , the  $J$ -th job required completing  $J(J + 1)/2 - (J - 1)J/2 = J$  additional real-effort tasks on the margin. Thus, subjects have to work more and longer to complete each subsequent job.<sup>8</sup> This further aligns the job choice with the canonical assumption of the cost being convex and the net utility being concave in jobs.

Such a feature should theoretically decrease the frequency of corner solutions, i.e. fewer participants should choose either the minimum or the maximum amount of tasks to complete. The reduction of corner choices makes it easier to estimate structural parameters based on the assumption of internal optima. The mapping

<sup>7</sup>This means that participants made  $5 \times 9 = 45$  decisions in total.

<sup>8</sup>For comparison, in Augenblick and Rabin (2019) and Chakraborty and Fenig (2022) the total earnings increase proportionally with effort.

from jobs to tasks was explained in advance, as well as transparently presented to the participants on every decision screen (see Figure 1).

### 2.2.2 All choices and predictions

Table 1 summarizes all the decisions made by participants by date.  $J_{t\tau}(w)$  denotes the decisions made on day  $t$  about the number of jobs to complete on day  $\tau$  at the wage  $w$ . On day 1 participants decided the number of jobs to complete on days 1, 2, and 3, denoted by  $J_{11}(w)$ ,  $J_{12}(w)$  and  $J_{13}(w)$ , respectively. On day 2, they decided the number of jobs to complete on days 2 and 3, i.e.  $J_{22}(w)$  and  $J_{23}(w)$ , respectively. On day 3, they decided how many jobs to complete on day 3, denoted by  $J_{33}(w)$ .  $J_{t\tau}^p(w)$  denotes the prediction of the decision  $J_{t\tau}(w)$  to be made on day  $t$  about the number of jobs to be done on day  $\tau$ . Participants faced 3 prediction screens. On day 1, they predicted  $J_{22}(w)$  and  $J_{23}(w)$ , i.e, the number of jobs that their day 2 self will choose on day 2 to complete on day 2 and day 3. And, on day 2, they predicted  $J_{33}(w)$ , the number of jobs that their day 3 selves will choose on day 3 to complete on day 3.

Table 1: Decisions by date

		Work date ( $\tau$ )		
		1	2	3
Decision date ( $t$ )	1	$J_{11}(w)$	$J_{12}(w), J_{22}^p(w)$	$J_{13}(w), J_{23}^p(w)$
	2		$J_{22}(w)$	$J_{23}(w), J_{33}^p(w)$
	3			$J_{33}(w)$

$J_{t\tau}(w)$  denotes the decision made on date  $t$  about the number of jobs to do on date  $\tau$ .  $J_{t\tau}^p(w)$  denotes the prediction about  $J_{t\tau}(w)$ .  $w \in \{1, 3, 5, 7, 9\}$  denotes the wage rate.

On each day  $\tau$ , one decision was randomly chosen from the pool of decisions made *for* that day, and participants had to complete the corresponding number of jobs for the corresponding wage. As an example, on day 3 one decision was randomly drawn from a pool of 15 decisions: decisions  $J_{13}$ ,  $J_{23}$ , and  $J_{33}$ , each of them made for 5 different wages.

## 2.3 Sessions

The sessions were run online in November 2021 and May 2022 with participants recruited from Prolific, an online participant recruitment platform. In November, the median length of the experiment was 30 minutes on the first day, 15 minutes on the second day, and 9 minutes on the third day. In May, the median length was 27 minutes, 12 minutes, and 12 minutes, respectively.

## 2.4 Payments

All payments were made within the evening of the third day after the survey was over.<sup>9</sup> There were four sources of payment for each participant. First, participants received a participation fee for each day they participated in the study (£2, £1 and £1, respectively). On the first day, the participation fee was higher than on the other two days as participants spent more time reading instructions, completing mandatory tasks for training, and making more decisions. Second, participants received a bonus of £5 if they completed all three days of the experiment. This completion fee offered was relatively high to minimize attrition.

Third, on each day participants completed tasks and were paid according to one randomly chosen decision out of the decisions they made about that day. To calculate this payment, the randomly drawn wage rate is multiplied by the number of jobs chosen and completed for that decision. On each day, the payment can range from 3p ( $1p \times 3$  jobs) to 117p ( $9p \times 13$  jobs).

Finally, participants could receive a fourth payment of 20p if one of their predictions (randomly chosen from the 15 predictions that participants made) differed by no more than two jobs from their actual decision.

**Incentives for prediction:** If the accuracy bonus for prediction is large, sufficiently sophisticated subjects can use the prediction tasks as a commitment device, as the bonus received from matching a previously-made prediction could incentivize their future selves to work harder than they would under just wage incentives. On the other hand, without any incentives, participants might not be

---

<sup>9</sup>After verifying a new participant's credentials and the first four cash outs, Prolific allows participants to instantly cash out their payments from all future studies. We received zero complaints about the payments not going through on time.

thoughtful about their prediction choices at all. As a compromise, we provided a very small accuracy bonus. Subjects completed fifteen prediction tasks at different wages and were paid at most 20p for one randomly chosen task: on average, every accurate prediction task only paid 1.33p. A strong argument in favor of incentivizing predictions also comes from [Augenblick and Rabin \(2019\)](#), who randomly assigned accuracy bonuses between \$0.25 and \$8.25 to identify if higher accuracy payments affected predictions or future decisions, and found no such evidence.

When participants made their actual decisions, we deliberately did not remind them of their predictions to avoid priming them into being consistent. To make sure that participants understood all the sources of payment, we quizzed them about the experiment's payment structure.

On the three survey days, participants earned 35p, 35p, and 32p on average for completing their chosen number of transcription tasks, respectively. Around 87% of all predictions were accurate within 2 jobs, and hence qualified for 20p reward if chosen for payment.

### 3 Model and Identification Strategy

The experimental design described in Section 2 allows us to estimate the present bias parameters over effort ( $\beta_e$ ) and money ( $\beta_m$ ), and the sophistication parameters over effort ( $\hat{\beta}_e$ ) and money ( $\hat{\beta}_m$ ). To estimate these parameters, we (implicitly) ask participants to trade off the "benefits of completing jobs" (i.e. the money  $Jw$ ) and the costs (i.e. the effort that must be exerted).

We assume that on decision day  $t_d$ , participants choose their optimal effort  $J^*$  for day  $t_e \geq t_d$  on the experimental task with piece-rate  $w$  based on the following utility optimization exercise:

$$J^* = \arg \max_J \beta_m^{\mathbb{1}(3 > t_d)} \delta^{3-t_d} Jw - \frac{1}{\psi} \beta_e^{\mathbb{1}(t_e > t_d)} \delta^{t_e-t_d} c(J_i(J_i + 1)/2). \quad (1)$$

The indicator  $\mathbb{1}(3 > t_d)$  denotes whether the payment day (day 3) is later than the decision day. Similarly,  $\mathbb{1}(t_e > t_d)$  indicates whether the decisions are being made about working in the future, and  $(t_e - t_d)$  measures the distance between

the effort date and the decision date.  $\psi$  is required as the benefits of exerting effort are measured in a different unit from the costs.

A slight variation of equation 1 also describes how a participant comes up with her predictions. In this variation,  $\beta_m$  is replaced with  $\hat{\beta}_m$  and  $\beta_e$  is replaced with  $\hat{\beta}_e$ .

In equation (1), the first term is the benefit of completing more jobs and the second term is its cost. The benefits are linearly increasing in jobs  $J$ , and hence the solution to equation (1) is interior only when the cost of effort is convex. The linearity in money and convexity of effort cost are standard identification assumptions used extensively in previous studies (Andreoni and Sprenger, 2012; Augenblick et al., 2015; Augenblick and Rabin, 2019). Further, these together guarantee that the net utility is concave in effort, which is necessary for an interior maximum.

By paying for jobs that convert to tasks (the unit of effort) through a convex transformation, we make it more likely that the subjects indeed face an increasing marginal cost of completing more jobs, and equivalently, the marginal utility of jobs is decreasing. For example, assuming a power function we get:  $c(J) = \frac{(J(J+1)/2)^\gamma}{\gamma}$ , where  $\gamma$  determines the convexity of the cost function. Even if transcription tasks require a constant marginal cost of effort and time ( $\gamma = 1$ ), the cost function is still convex in  $J$ .

Overall, as long as  $c(\cdot)$  is convex in  $J$ , the problem in equation (1) is strictly concave and has a unique solution.

$$\left( \frac{\beta_m^{\mathbb{1}(3 > t_d)} w \psi \delta^{3-t_e}}{\beta_e^{\mathbb{1}(t > t_d)}} \right) - \left( \frac{J(J+1)}{2} \right)^{\gamma-1} \left( \frac{2J+1}{2} \right) = 0. \quad (2)$$

As before, to get the equivalent condition for predictions,  $\beta_m$  is replaced with  $\hat{\beta}_m$ , and  $\beta_e$  is replaced with  $\hat{\beta}_e$ . To estimate the parameters of the model, i.e.  $\theta = (\beta_e, \hat{\beta}_e, \beta_m, \hat{\beta}_m, \delta, \psi, \gamma)$ , we use the maximum likelihood estimator (MLE) over both choice and prediction data.

Our estimate,  $\hat{\theta}_{\text{MLE}}$ , is defined as

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta} \sum_{i=1}^N \ln \mathcal{L}_i(\theta), \quad \text{where}$$

$$\ln \mathcal{L}_i(\theta) = -\ln \sigma - 0.5 \ln 2\pi - 0.5 \left( \frac{J_i^*(\theta) - J_i}{\sigma} \right)^2. \quad (3)$$

$N$  denotes the number of data points and  $J_i^*(\theta) - J_i = \varepsilon_i$  is the error term, which is assumed to follow a standard distribution with standard deviation  $\sigma$ . Any combination of preference parameters in (2) implies a unique  $J_i$ . MLE finds the set of parameters that maximizes the likelihood of observing the collected data.

### 3.1 How conflated estimates might be biased:

Suppose, the observed choice and prediction data is generated under true preferences of  $1 = \hat{\beta}_m = \beta_m > \hat{\beta}_e > \beta_e$  and some  $\delta \leq 1$ , as previously suggested in the literature (and later confirmed by our data), but the analyst assumes  $\beta_m = \beta_e$  and  $\hat{\beta}_m = \hat{\beta}_e$ .<sup>10</sup> How would that *restricted model* change or bias the estimates the analyst would obtain?

We start with the true parameters  $1 = \hat{\beta}_m = \beta_m > \hat{\beta}_e > \beta_e$  and intuitively argue how the MLE would re-adjust those values when estimating a restricted model. First, fix the decisions where both the effort and payment are in the future. In this case, the first term of the structural equation (2) has a  $\frac{\beta_m}{\beta_e}$  expression in it. When the restricted model enforces  $\beta_m = \beta_e$ , the  $\frac{\beta_m}{\beta_e}$  expression decreases to 1, and the MLE would try to compensate for that by increasing the  $\delta$  in the numerator.<sup>11</sup> The change need not be one-to-one, but this would be the qualitative direction of the change. Further, for a fixed original  $\frac{\beta_m}{\beta_e}$ , the smaller the time delay  $D$  between effort and payment ( $D = 3 - t_e$  days here), the more  $\delta$  has to increase to compensate for  $\frac{\beta_m}{\beta_e}$ . For example, if  $\delta$  increases by a factor of 1.03 under  $D = 7$ , it would have to increase by a factor of  $(1.03)^7 = 1.23$  under a smaller  $D = 1$ . Thus, in a 3-day

<sup>10</sup>We assume that similar to our data, this dataset has adequate variation in the delay to effort and the delay to payment.

<sup>11</sup>The value of  $\psi$  is pinned by the decisions where both the payment and effort are immediate. For these, neither  $\beta$  nor  $\delta$  play any role.

design like ours, where  $D$  is small, the effect of misspecification on the upwards bias of  $\delta$  will be especially prominent.

Next, fix decisions where the payment is in the future and the effort is *immediate*. In equation (2), the  $\beta$  disappears from the denominator and only appears on the numerator. To compensate for the higher  $\delta$  in the numerator as discussed by the previous point, the common  $\beta$  has to be low (and thus resemble  $\beta_e$ ) rather be high (i.e., resemble  $\beta_m$ ).

Finally, given the assumption on the data-generating process, predictions made about choices with future payments would be accurate as  $\beta_m = \widehat{\beta}_m$ , whereas those made about choices with future effort would fall short of reality as  $\beta_e < \widehat{\beta}_e$ . As, the conflated  $\beta$  and  $\widehat{\beta}$  would have to account for both these effects, the relative sophistication would be higher than that for effort.

## 4 Results

### 4.1 Monotonicity and Corner Choices

Before our core analysis, we check for internal consistency in the data through the frequency of monotonicity violations with respect to money and time delay (Chakraborty et al., 2017).

**Wage Monotonicity:** For the data to be consistent with any rational model of choice, participants should decide to complete more jobs for higher wages. As described in Section 2, on each of the 9 decision screens, participants choose (or predict) the number of real-effort tasks to complete at 5 different wage rates for a fixed work-date  $\tau$  and decision-date  $t$ . Wage-monotonicity (weak) is violated if for some fixed decision-screen  $(t, \tau)$ , participants choose to (or predict to) complete strictly fewer jobs for a higher wage rate, i.e,  $J_{t\tau}(w) < J_{t\tau}(w')$  when  $w > w'$ . For each of the 108 participants who completed all 3 days of the study, we measured the number of decision screens with wage-monotonicity violations. As shown in Table 2, the majority of participants do not exhibit any wage-monotonicity violations. The number of participants violating strict wage-monotonicity on more than one out of the 9 decision screens is about 16 percent of the sample.

	Number screens where monotonicity is violated										Total
	0	1	2	3	4	5	6	7	8	9	
Number of participants	58	33	5	1	2	2	4	1	1	1	108

Table 2: Frequency of monotonicity violations (incorrect in brackets)

**Time Monotonicity:** One can also check for time-monotonicity (with respect to effort date), which is violated if, for a fixed wage rate and decision date, a participant decided to or predicted to work less on a later day  $\tau + 1$  than an earlier day  $\tau$ , i.e.,  $J_{t\tau}(w) < J_{t(\tau+1)}(w)$ . There are 20 such pairwise comparisons<sup>12</sup> per individual and we find only 15.37% cases with violations.

A second violation of time-monotonicity (with respect to pay date) happens when  $J_{t\tau}(w) > J_{(t+1)(\tau+1)}(w)$ , as the payment period is closer to the decision day on the right-hand side, whereas, the effort period is equidistant ( $\tau - t$  days) from the decision day on both sides. There are 15 such pair-wise comparisons per participant, and we find 28.83% cases with violations.

**Corner choices:** Out of all 4,860 choices made by the 108 participants, 1,207 choices (24.8 percent) are at the lower corner as participants chose the minimum number of jobs to complete, and 644 choices (13.3 percent) are at the upper corner as participants chose the maximum number of jobs to complete. For comparison, we get significantly fewer corner choices than the CTB method as implemented by [Andreoni and Sprenger \(2012\)](#) (70% corner choices ).

## 4.2 Selection

Out of the 137 participants who completed the first day of the study, 108 responded to all the questions on 3 days. Out of these 108 participants, 7 participants have no variation in any of the 9 screens, leaving us with 101 participants on whom we can run a regression meaningfully. Out of these 101 participants, we also remove the participants who ranked in the top 10% of "most frequent violators" in each category of wage monotonicity, time-monotonicity with respect to effort date, and

<sup>12</sup>One can create more comparisons using transitivity, by also comparing  $\tau + 2$  to  $\tau$ , but that runs into double-accounting of violations.



time-monotonicity with respect to pay date. Our final data consists of 80 (out of 108) participants.

To remain consistent throughout our results, all the tables or graphs that follow use the data from all these 80 participants. In Appendix 6.1, we re-estimate our main aggregate structural results while only dropping the participants who were in the top 5% of violations in any of the three categories listed above, and we get very similar results with that sample of 90 participants.

### 4.3 Nonparametric Analysis

We start by showing the change in the participants' effort choices with respect to all the experimental variations. Following [Augenblick and Rabin \(2019\)](#), we employ the following specification at the aggregate level,

$$e_{i,j} = \phi_0 + \phi_e^{\text{future}} 1_{i,j}^{\text{future } e} + \phi_e^{\text{predict}} 1_{i,j}^{\text{predict } e} + \phi_m^{\text{future}} 1_{i,j}^{\text{future } m} + \phi_m^{\text{predict}} 1_{i,j}^{\text{predict } m} \quad (4) \\ + \phi_\gamma w_{i,j} + \phi_\delta (3 - t_e) + \mu_i + \varepsilon_{i,j}.$$

where  $e_{ij}$  is the number of jobs (or tasks) chosen by individual  $i$  at choice  $j$ . The regression controls for standard impatience or discounting through  $(3 - t_e)$ , the relative distance between the payment date and the effort date.<sup>13</sup> To allow for different base levels of efforts chosen by different individuals, we include participant-level fixed effect terms  $\mu_i$  as dummy variables. We also control for the wage rate  $w_{i,j}$ .

$\phi_e^{\text{future}}$  is multiplied with the indicator that takes a value of 1 only if the observation was an effort choice for a future date:  $\phi_e^{\text{future}}$  measures the additional number of jobs that participants want to complete in the future, as compared to the present.<sup>14</sup> If  $\phi_e^{\text{future}}$  is close to zero, an individual is not present biased, as the chosen effort levels are independent of the timing of the task. The larger is  $\phi_e^{\text{future}} > 0$ , the more an individual is present-biased (smaller  $\beta_e$ ), as, ceteris paribus, she would rather work more in the future than in the present. Thus,  $-\phi_e^{\text{future}}$  provides a measure of (and is conceptually rank-correlated with)  $\hat{\beta}_e$ . Similarly,  $-\phi_e^{\text{predict}}$  provides

<sup>13</sup>Equation 2 shows why long-run discounting only affects effort choice through  $(3 - t_e)$ .

<sup>14</sup>Note that this interpretation is possible as we already control for standard discounting.

a measure of  $\hat{\beta}_e$ , the sophistication over effort.

Similarly,  $\phi_m^{\text{future}}$  is multiplied with an indicator that takes a value of 1 only if the payment day was the future. The lower  $\beta_m$  is, the lower (and negative)  $\phi_m^{\text{future}}$  should be. Thus  $\phi_m^{\text{future}}$  provides a measure of (and is conceptually rank-correlated with)  $\beta_m$ . Similarly,  $\phi_m^{\text{predict}}$  provides a measure of sophistication over money.

### 4.3.1 Aggregate Analysis

Table 3 shows non-parametric measures of present bias and sophistication over the effort and money domains. As our dependent variable, we use both the number of jobs chosen, which is directly related to the payment (wage) that participants earn, and the number of real-effort tasks.

Table 3: Nonparametric estimates of present bias and sophistication at the aggregate level

	Jobs	Tasks
<i>Panel A: Effort domain</i>		
$\phi_e^{\text{future}}$	0.52*** (0.13)	4.17*** (1.11)
$\phi_e^{\text{predict}}$	0.33** (0.15)	3.09** (1.32)
<i>Panel B: Money domain</i>		
$\phi_m^{\text{future}}$	-0.05 (0.16)	-0.52 (1.34)
$\phi_m^{\text{predict}}$	-0.07 (0.14)	-0.56 (1.21)
<i>Panel C: Other coefficients</i>		
$\phi_\gamma$	0.42*** (0.01)	3.27*** (0.09)
$\phi_\delta$	0.12 (0.08)	1.23* (0.73)
N	80	80

Fixed effects regression estimates of Equation (4). \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ . The dependent variables are the number of Jobs or Tasks chosen.

Participants choose .52 more jobs or 4.17 more transcription tasks (which translates to 14% more tasks<sup>15</sup>) to be completed in the future than in the present, but

<sup>15</sup>The percentage is calculated by dividing the raw number by 29.39, the average number of tasks participants completed.

believe that this effect would be smaller (.33 jobs or 3.09 tasks), which is evidence of partial sophistication. The difference between  $\phi_e^{\text{future}}$  and  $\phi_e^{\text{predict}}$  is not statistically significant. We do not observe a statistically significant degree of present bias in the money domain. Participants choose .05 fewer jobs or 0.52 fewer transcription tasks (which translates to 1% fewer tasks) when they are paid on the same day than when they are paid in the future. Neither of the quantities is significantly different from zero. This result is consistent with findings by [Andreoni and Sprenger \(2012\)](#) and [Augenblick et al. \(2015\)](#), who do not find evidence of present bias over money. We also do not find a significant degree of sophistication over money.

#### 4.3.2 Individual Analysis

To derive non-parametric measures of present bias and sophistication for each participant, we use the regression in Equation (4) at the individual level for all 80 participants. Below, we only report the results using the number of jobs, as the results are qualitatively identical if we used the number of tasks instead.

Figure 3 shows the histogram of nonparametric estimates of present bias and sophistication in both the effort and the money domains, for all but one participant, who has an outlier estimate of sophistication over effort of 12.6. The top-left panel shows the estimates of present bias in the effort domain. The  $\phi_e^{\text{future}}$  estimates of 47 out of 79 participants are below 0, which indicates a present bias over effort. The median is  $-0.2$  and the mean is around  $-0.53$ . The top-right panel shows that most participants are at least partially aware of being present biased over effort. The median and mean of  $\phi_e^{\text{predict}}$  are  $-0.14$  and  $-0.17$ , respectively. The magnitude of this estimate is significantly smaller than that of present bias, which indicates that participants are indeed only partially sophisticated. In the money domain, the median and mean of the present bias estimate are  $-0.06$  and  $-0.03$ , respectively. Again, participants appear to be partially sophisticated.

Figure 4 uses 4 scatter plots to show the relationship between present bias and sophistication across effort and money domains. Individual estimates under zero indicate the presence of present bias or sophistication along the respective axes. We find a strong positive relationship between present bias over effort and over money (top-left panel, Spearman coefficient= $0.75$ ,  $p < 0.0001$ ), or between present

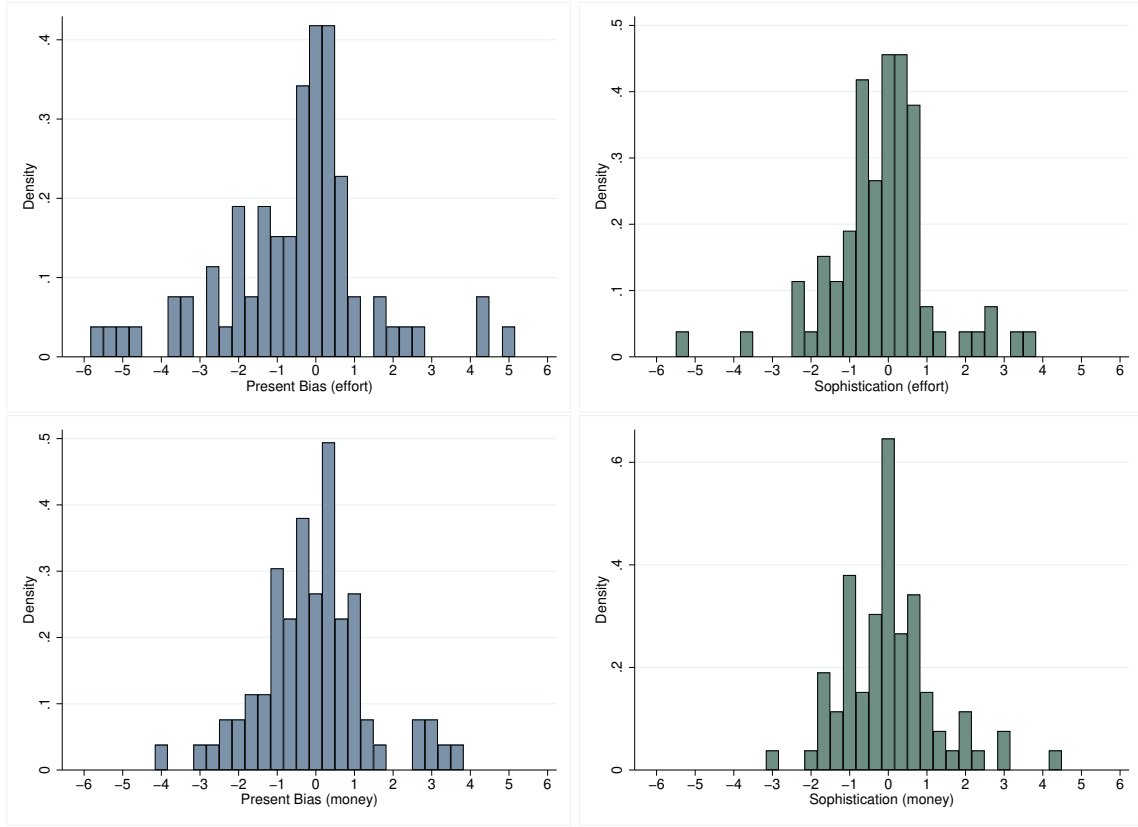


Figure 3: Histogram of non-parametric estimates of present bias and sophistication, estimated at the individual level. *Present bias* and *Sophistication* for effort corresponds to  $-\phi_e^{\text{future}}$  and  $-\phi_e^{\text{predict}}$ , respectively. *Present bias* and *Sophistication* for money correspond to  $\phi_m^{\text{future}}$  and  $-\phi_m^{\text{predict}}$ , respectively.

bias and sophistication over effort (bottom-left panel, Spearman coefficient=0.72,  $p < 0.0001$ ). This indicates that individuals who are present biased over effort tend to also be sophisticated about that present bias and are also present biased over money. We also observe a positive, but less strong relationship between sophistication over effort and over money (bottom-right panel, Spearman coefficient=0.34,  $p = 0.0019$ ).

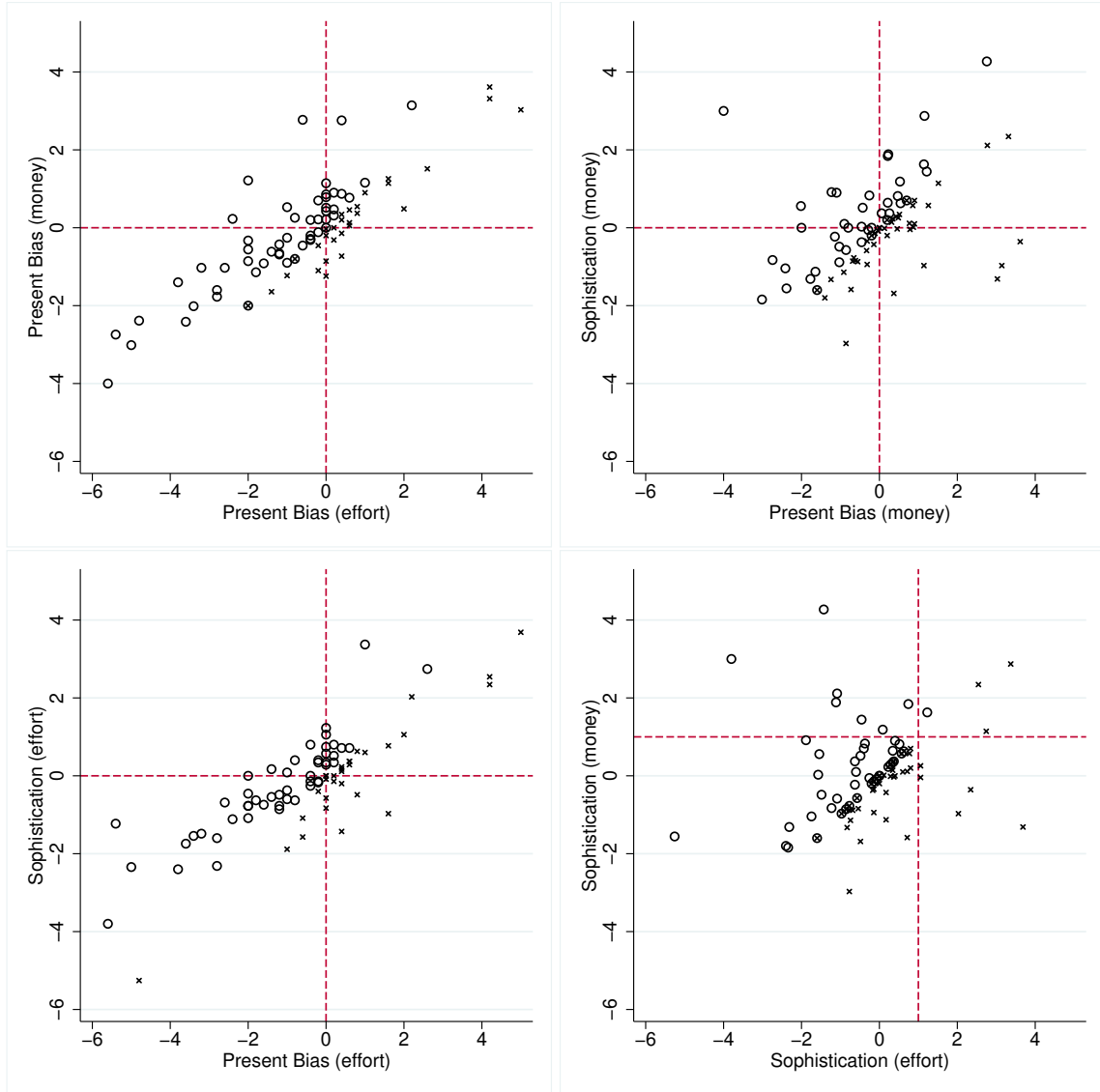


Figure 4: Comparison of nonparametric estimates of present bias and sophistication across domains. To make the comparisons easier, observations above and below the 45-degree line are marked by circle and cross marks respectively.

## 4.4 Structural Estimation

### 4.4.1 Visualizing the structural identification:

Before our structural estimation of aggregate data, we plot certain features of the aggregate data that provide an intuition about how the parameters of the structural

model are identified. For simplicity, we assume  $\delta = 1$ .<sup>16</sup> Thus, any discounting must come from the present bias  $\beta_e$  and  $\beta_m$ . Let  $e_{t\tau}$  be the effort chosen on day  $t$  for day  $\tau$ .

For a fixed decision period  $s$  and payment period, the comparison of immediate choice ( $e_{ss}$ ) to future choice ( $e_{st}$ ) reveals  $\beta_e$ . Under  $\beta_e \leq 1$ , we would observe  $e_{ss} < e_{st}$  for all  $s < t$ , and under  $\beta_e = 1$  observe  $e_{ss} = e_{st}$ . In the right panel of Figure 5, we compare the empirical CDFs of immediate effort  $e_{ss}$ , as measured by tasks<sup>17</sup>, and future effort  $e_{st}$  while pooling over wage rates and decision days ( $ss = 11, 22, 33$  and  $st = 12, 23$ ).<sup>18</sup>  $e_{ss}$  is first-order stochastically dominated by  $e_{st}$ , hinting that fewer tasks are allocated to the immediate day consistently. A Kolmogorov-Smirnov test finds that the two empirical distributions are significantly different with  $p = .001$ . This suggests that the structural estimation could yield  $\beta_e < 1$ .

Sophistication about  $\beta_e$  is identified if the participants predict the difference between immediate and future choices. In particular, under  $\hat{\beta}_e \leq 1$ , we would observe  $\hat{e}_{22} < \hat{e}_{23}$ , and under  $\hat{\beta}_e = 1$  observe  $\hat{e}_{22} = \hat{e}_{23}$ . The present bias for money is identified by comparing two immediate choices  $e_{33}$ ,  $e_{22}$ , where  $e_{33}$  pays on the same day and hence is not affected by  $\beta_m$ . Under  $\beta_m < 1$ , we would observe  $e_{33} > e_{22}$  and under  $\beta_m = 1$  we would observe  $e_{33} = e_{22}$ . Finally, the prediction of immediate effort  $\hat{e}_{22}$  is not affected by  $\beta_e$  but is affected by  $\hat{\beta}_m$ . Thus, under  $\hat{\beta}_m < 1$ , we would observe  $e_{33} > \hat{e}_{22}$  and under  $\hat{\beta}_m = 1$  we would observe  $e_{33} = \hat{e}_{22}$ .

To get an idea about  $\hat{\beta}_e$ ,  $\beta_m$ ,  $\hat{\beta}_m$ , we plot the empirical CDFs of  $e_{22}$ ,  $\hat{e}_{23}$ ,  $e_{33}$ ,  $\hat{e}_{22}$  in the right panel of Figure 5. The empirical CDFs are all close to each other and intertwined, with none of the distributions first-order stochastically dominating each other. Similarly, none of three pairwise comparisons,  $\hat{e}_{22}$  vs  $\hat{e}_{23}$ ,  $e_{33}$  vs  $e_{22}$ ,  $e_{33}$  vs  $\hat{e}_{22}$  (the difference between which identify  $\hat{\beta}_e, \beta_m, \hat{\beta}_m < 1$  respectively) are significantly different under the Kolmogorov-Smirnov test. Thus based on these results, ex-ante we do not expect to estimate  $\hat{\beta}_e, \beta_m, \hat{\beta}_m$  values that are different than 1.

<sup>16</sup>Later, our structural estimates will confirm that it is a fair approximation for our data.

<sup>17</sup>Using number of jobs leads to identical conclusions.

<sup>18</sup>Assuming  $\delta = 1$  allows us to also include  $e_{33}$  for the immediate versus future comparison. Not including it would still lead to identical conclusions.

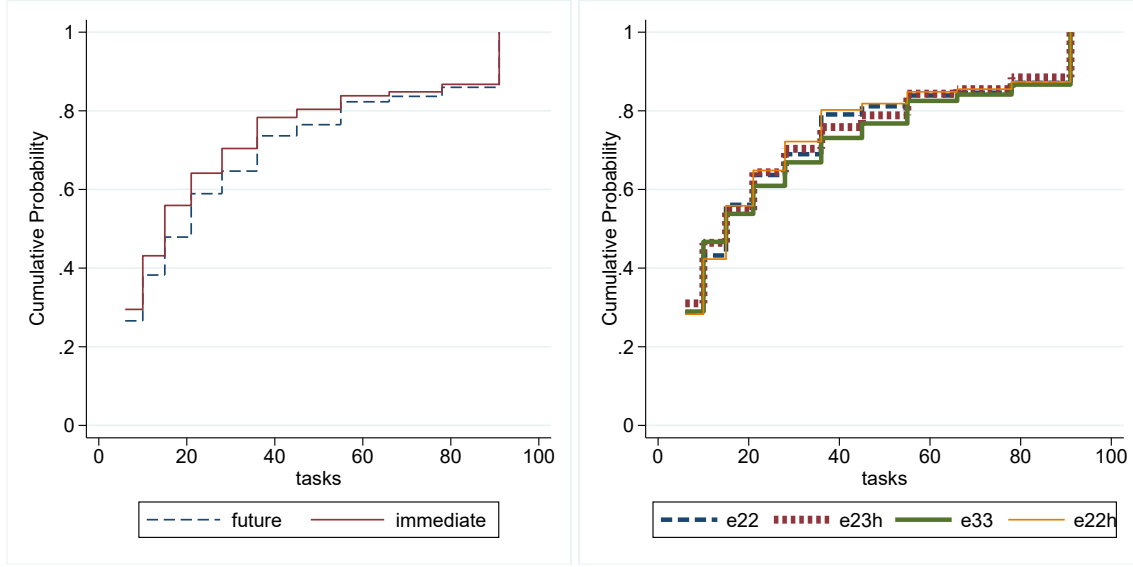


Figure 5: The left figure plots immediate versus future effort, and helps identify  $\beta_e$ . The right figure plots  $e_{22}$ ,  $\widehat{e}_{23}$ ,  $e_{33}$ ,  $\widehat{e}_{22}$ , and helps identify  $\widehat{\beta}_e, \beta_m, \widehat{\beta}_m$ .

Next, we structurally estimate the parameters  $\beta_e$ ,  $\beta_m$ ,  $\widehat{\beta}_e$ , and  $\widehat{\beta}_m$ , both at the aggregate level and at the individual level. Both of our analyses use the quasi-hyperbolic discounting model described in equation (1) and the maximum likelihood estimator described in Section 3.

#### 4.4.2 Aggregate Analysis

Table 5 presents the structural estimates at the aggregate level. For each parameter, we report the 95% confidence interval instead of the standard error, to make it easy to understand if the coefficients are significantly different than 1 or not.

In the first column (“Same”), following [Augenblick and Rabin \(2019\)](#), we assume that present bias and sophistication are *identical* across money and effort. Thus, this specification replicates their identification assumption and serves as a counterfactual. The estimated present bias parameter is 0.71 and the 95% confidence interval lies below 1. The sophistication parameter is 0.74 and again, the 95% confidence interval lies below 1.

The second column (“Separate”) is our novel specification: it allows for different present bias and sophistication over effort and money. We find that  $\beta_e$  is 0.70 and is significantly different from 1. Interestingly, the common  $\beta$  estimated in

the “Same” specification mimics the  $\beta_e$  in the “Separate” specification.  $\hat{\beta}_e$  is 0.80, which is higher than  $\beta_e = .70$  (though the difference is not statistically significant) and indicates partial sophistication over effort. Neither of the money-domain parameters  $\hat{\beta}_m = \beta_m = .95$  are estimated to be significantly different from 1.

Researchers interested in estimating only the present bias parameters could use a variant of our design that does not include the prediction tasks. To approximate the results of a design with no predictions, in the third column (“No prediction”) we run our analysis while excluding the prediction tasks (i.e., using only the data from the 6 decision screens that are not predictions). The results are very similar to the results from the “Separate” specification. This is reassuring, as it indicates that including the prediction tasks does not bias the estimates of present bias.

The 95% confidence interval of  $\delta$  contains 1 consistently across all three specifications. Motivated by this finding, in the fourth column (“Separate,  $\delta = 1$ ”), we re-estimate the *Separate* specification assuming  $\delta$  to be equal to 1 and including both choice and prediction data. While this increases the estimates of present bias over both domains,  $\beta_e$  is still estimated to be significantly below 1.

### **Model comparison by Likelihood Ratio (LR) test of nested models & AIC**

Consider the null hypothesis that the simpler nested model in specification [1] fits the data as well as the model in specification [2]. The LR test statistic in the comparison of [1] and [2] is 10.92, which should be distributed as chi-squared with 2 degrees of freedom. We obtain a corresponding p-value  $< .01$  and thus we have statistically significant evidence to reject the null and conclude that [2] fits the data better than [1]. On the other hand, the LR test statistic in the comparison of [2] and [4] is 1, which should be distributed as chi-squared with 1 degree of freedom. The corresponding p-value is .68 and we do not find significant evidence for model [2] fitting the aggregate data better than the simpler nested model [4]. Thus [4] fits the data almost as well as any other model without overfitting the data. Akaike Information Criterion (AIC) for model selection also suggests [4] as the superior model among [1], [2] and [4].

### **Stability**

All estimated parameters are relatively stable across all specifications. For exam-



ple, under all specifications,  $\beta_e$  is significantly below 1, and  $\hat{\beta}_e, \beta_m$  are no different than 1. Similarly, the convexity parameter  $\gamma$  is consistently above one, implying that the cost function is significantly convex.<sup>19</sup> The scale of effort  $\psi$  is also similar across all specifications, implying that the baseline level of effort did not vary with the type of model chosen.

#### Explaining conflated estimates from Model [1]:

The conflated estimates confirm our hypotheses from Section 3.1. In model [1], MLE tries to fit the misspecified model to the data by biasing  $\delta$  upwards (increasing  $\delta$  significantly over 1). The common  $\beta$  is closer to  $\beta_e$  than to  $\beta_m$ . Further, the estimated  $\beta$  and  $\hat{\beta}$  are almost identical in magnitude under [1], implying a high degree of estimated relative sophistication. In fact, [1] is the only specification that rejects complete naivete ( $\hat{\beta} = 1$ ) in favor of  $\hat{\beta} < 1$ . In this way, subjects are estimated as relatively more sophisticated under [1] than in all other specifications.

#### 4.4.3 Individual Analysis

Next, we estimate the time preference parameters for each individual separately using the maximum likelihood estimator based on Equation 3. The results are best illustrated graphically. Figure 6 plots the distributions of the present bias and sophistication parameters, in both the effort and the money domains, for 77 participants.<sup>20</sup> The top-left and bottom-left panels present the present bias parameters over effort and over money, respectively. The estimated parameter for present bias over effort,  $\beta_e$ , is below 1 for over 67 percent of the sample. For four participants,  $\beta_e$  is less than 0.1.<sup>21</sup> The median and mean for present bias over effort are 0.91 and 0.98, respectively. The median and mean of  $\beta_m$  are 1.04 and 1.28, respectively. This confirms that present bias over money is weaker than present bias over effort.

The top-right and bottom-right panels present the sophistication parameters

---

<sup>19</sup>The convexity of the cost function guarantees that the problem in equation (1) allows for a unique solution if we have an interior solution.

<sup>20</sup>To make the histogram easy to read, we have removed three outliers with estimates of present bias or sophistication above 16.

<sup>21</sup> $\beta_e$  is close to 0 for participants who (i) decide to complete many more tasks in the future compared to the present and (ii) do not react strongly to changes in wage rates.

Table 4: Structural estimates at the aggregate level

	[1] Same (AR2019)	[2] Separate	[3] No prediction	[4] Separate, $\delta = 1$
Present bias ( $\beta$ )	0.71 (0.52,0.91)			
Sophistication ( $\hat{\beta}$ )	0.74 (0.56,0.93)			
Present bias, effort ( $\beta_e$ )		0.70 (0.51,0.90)	0.71 (0.52,0.91)	0.79 (0.65,0.93)
Present bias, money ( $\beta_m$ )		0.95 (0.64,1.26)	0.93 (0.60,1.27)	1.09 (0.88,1.30)
Sophistication, effort ( $\hat{\beta}_e$ )		0.80 (0.53,1.07)		0.88 (0.63,1.13)
Sophistication, money ( $\hat{\beta}_m$ )		0.95 (0.66,1.23)		1.04 (0.79,1.29)
Discount factor $\delta$	1.19 (1.00,1.39)	1.10 (0.90,1.29)	1.09 (0.90,1.28)	
Convexity $\gamma$	2.56 (2.38,2.66)	2.58 (2.40,2.77)	2.58 (2.29,2.87)	2.57 (2.39,2.75)
Scale of effort $\psi$	199.36 (108.02,290.70)	208.41 (84.17,332.65)	211.52 (8.75,414.29)	199.40 (86.86,311.93)
Size of error $\sigma$	3.23 (3.17,3.30)	3.23 (3.16,3.30)	3.22 (3.15,3.30)	3.23 (3.16,3.29)
$\ln \mathcal{L}$	-6026.20	-6020.74	-4010.02	-6021.24
$N$	80	80	80	80
Observations	3600	3600	2400	3600

[1] *Same* follows [Augenblick and Rabin \(2019\)](#) to assume that present bias or sophistication over effort and money are identical. [2] *Separate* generalizes *Same* by allowing for different present bias and sophistication parameters across effort and money domains. [3] *No prediction* is similar to *Separate*, but excludes predictions and hence cannot estimate sophistication. [4] *Separate*,  $\delta = 1$  imposes that the long-run discount factor is one. The numbers in parentheses indicate the lower and upper bound of the 95% confidence interval.

over effort and over money, respectively. The median and the mean of the parameter for sophistication in the effort domain,  $\hat{\beta}_e$ , are 0.94 and 1.07, respectively. The median and the mean of the parameter for sophistication in the money domain,  $\hat{\beta}_m$ , are 1.05 and 1.29, respectively.

Figure 7 presents four scatterplots of the estimated individual-level present bias and sophistication parameters in both domains. The top-left panel presents the scatter plot of  $\beta_e$  and  $\hat{\beta}_e$ . The correlation between the two parameters is very high (coefficient = 0.96,  $p < .001$ ) and much higher than the range 0.24-0.28 reported in [Augenblick and Rabin \(2019\)](#) as the correlation coefficient between  $\beta$  and  $\hat{\beta}$ . The correlation suggests that participants are at least partially aware of their own

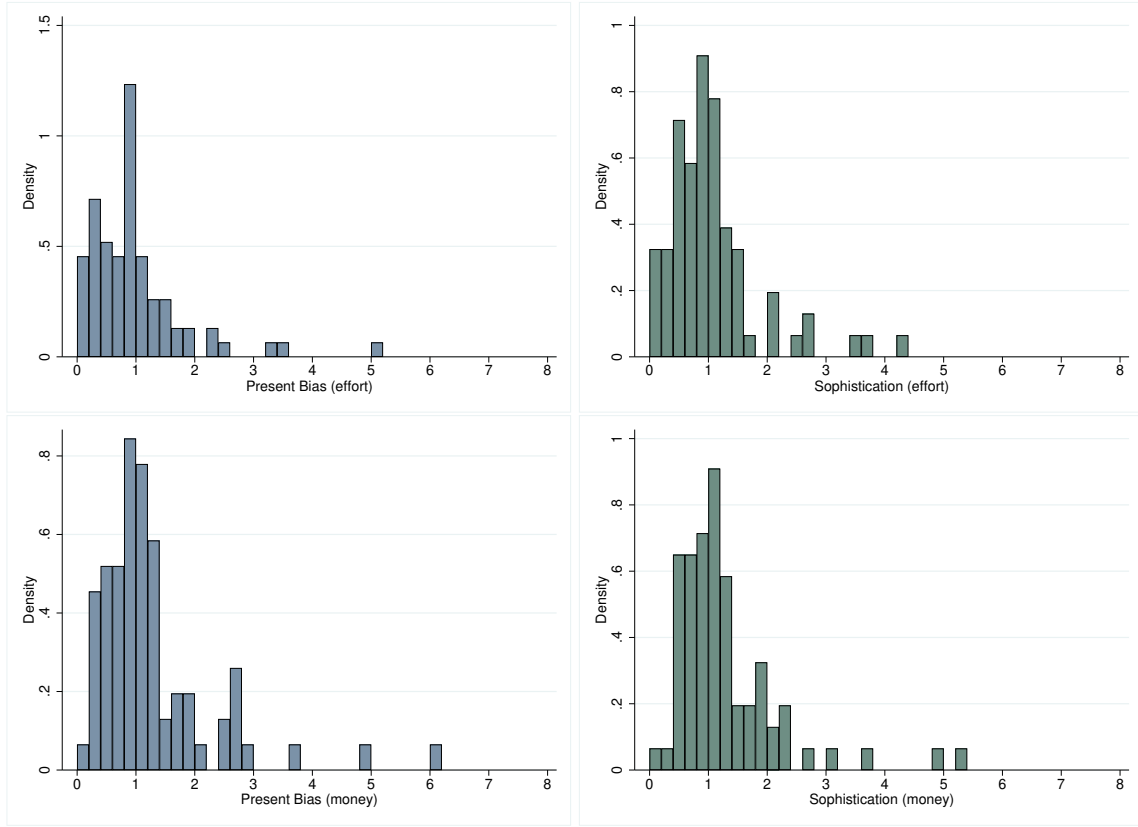


Figure 6: Structural estimates of present bias and sophistication over effort and money.

present bias over effort. The gulf in our and [Augenblick and Rabin \(2019\)](#)'s results could be explained by how their estimate of correlation also depends on the correlation between  $\beta_m$  and  $\hat{\beta}_m$ . In our data,  $\beta_m$  and  $\hat{\beta}_m$  have a correlation coefficient of  $-.06$ , which is not significantly different from zero. The top-right panel presents the scatter plot between  $\beta_m$  and  $\hat{\beta}_m$ .

The bottom-left panel in Figure 7 shows a strong and positive relationship between  $\beta_e$  and  $\beta_m$  (correlation  $0.88$ ,  $p$  value  $< .001$ ). The bottom-right panel shows no correlation ( $-0.04$ ,  $p$  value  $= .72$ ) between  $\hat{\beta}_e$  and  $\hat{\beta}_m$ .

Further, the parametric estimates of  $\beta_e$  ( $\hat{\beta}_e$ , respectively) are strongly correlated with the non-parametric estimates of present bias (sophistication, respectively) over effort obtained in Section 4.3.2, with a correlation coefficient of  $.45$  ( $.37$ , respectively). This correlation provides in-sample validation of the structural parameter estimates. All these correlations are statistically significant with  $p < .01$ .

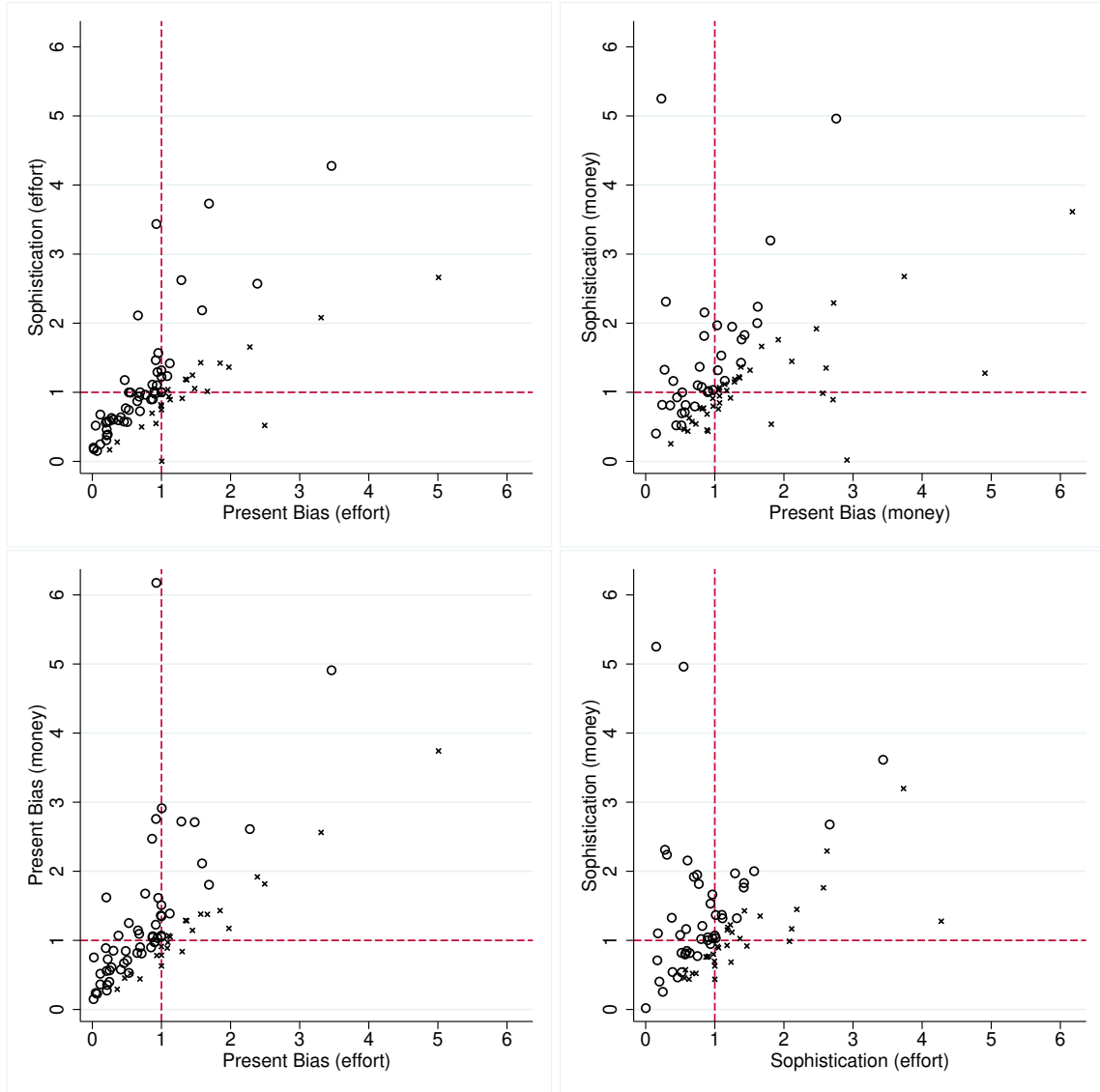


Figure 7: Comparison of structural estimates of present bias and sophistication across domains. The points above and below the 45-degree line are marked with a circle and check mark respectively.

## 5 Discussion and Conclusion

In this paper, we developed and implemented a novel experimental design to jointly estimate present bias and sophistication. Elicitation studies generally estimate discounting and present bias by asking participants to choose between time-dated monetary payments ([Andersen et al., 2008](#); [Andreoni and Sprenger, 2012](#);

Andersen et al., 2014; Andreoni et al., 2015; Halevy, 2015) or between time-dated effort tasks (Augenblick et al., 2015).<sup>22</sup> We contribute to this literature by additionally eliciting sophistication.<sup>23</sup>

Our design has three key and novel features. First, we vary both the delay of task completion and the delay of the payment. This allows us to estimate present bias and sophistication *separately* for the effort domain and for the money domain. Second, under our novel incentive structure, the number of corner solutions is substantially lower than that in comparable studies, which facilitates our structural estimation exercise. Third, unlike previous studies that lasted over multiple weeks, we implemented our experiment over three consecutive days. This makes our design cheaper and easier to implement, less prone to attrition, and more easily replicable for researchers who want to measure present bias and sophistication as control variables in a larger study.

Our non-parametric aggregate analysis shows that on average participants allocate 4 more tasks to the future than to the present, but predict this effect to be only 3 more tasks, thus showcasing partial sophistication. We find significantly weaker evidence for present bias or sophistication over money. Our structural analysis replicates these findings: consistent with Augenblick et al. (2015), we find significantly more present bias over effort than over money. The findings justify the use of a design that treats present bias/ sophistication over money and present bias/ sophistication over effort differently.

---

<sup>22</sup>Experiments have also elicited intertemporal choices over irritating noises (Solnick et al., 1980), squirts of juice (McClure et al., 2007), or snack foods (Read and Van Leeuwen, 1998).

<sup>23</sup>Freeman (2021) provides an alternative theoretical framework for revealing preference and sophistication types from behavior in a general environment that includes that of O'Donoghue and Rabin (1999).

## References

- S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutström. Eliciting Risk and Time Preferences. *Econometrica*, 76(3):583–618, 2008. ISSN 0012-9682. URL <https://www.jstor.org/stable/40056458>.
- S. Andersen, G. W. Harrison, M. I. Lau, and E. E. Rutström. Discounting behavior: A reconsideration. *European Economic Review*, 71:15–33, Oct. 2014. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2014.06.009. URL <https://www.sciencedirect.com/science/article/pii/S0014292114000968>.
- J. Andreoni and C. Sprenger. Estimating Time Preferences from Convex Budgets. *American Economic Review*, 102(7):3333–3356, Dec. 2012. ISSN 0002-8282. doi: 10.1257/aer.102.7.3333. URL <https://www.aeaweb.org/articles?id=10.1257/aer.102.7.3333>.
- J. Andreoni, M. A. Kuhn, and C. Sprenger. Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior & Organization*, 116:451–464, Aug. 2015. ISSN 0167-2681. doi: 10.1016/j.jebo.2015.05.018. URL <https://www.sciencedirect.com/science/article/pii/S0167268115001535>.
- J. Andreoni, C. Gravert, M. A. Kuhn, S. Saccardo, and Y. Yang. Arbitrage Or Narrow Bracketing? On Using Money to Measure Intertemporal Preferences. Working Paper 25232, National Bureau of Economic Research, Nov. 2018. URL <http://www.nber.org/papers/w25232>. Series: Working Paper Series.
- D. Ariely and K. Wertenbroch. Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science*, 13(3):219–224, 2002.
- N. Ashraf, D. Karlan, and W. Yin. Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics*, 121(2):635–672, 2006.
- N. Augenblick and M. Rabin. An Experiment on Time Preference and Misprediction in Unpleasant Tasks. *Review of Economic Studies*, 86(3):941–975, May 2019.

- ISSN 0034-6527. doi: 10.1093/restud/rdy019. URL <https://academic.oup.com/restud/article/86/3/941/4996235>.
- N. Augenblick, M. Niederle, and C. Sprenger. Working over Time: Dynamic Inconsistency in Real Effort Tasks. *The Quarterly Journal of Economics*, 130(3):1067–1115, Aug. 2015. ISSN 0033-5533. doi: 10.1093/qje/qjv020. URL <https://academic.oup.com/qje/article/130/3/1067/1934963>. Publisher: Oxford Academic.
- A. Chakraborty. Present bias. *Econometrica*, 89(4):1921–1961, 2021.
- A. Chakraborty and G. Fenig. From me to we: Beating procrastination in teams. 2022.
- A. Chakraborty, E. M. Calford, G. Fenig, and Y. Halevy. External and internal consistency of choices made in convex time budgets. *Experimental Economics*, 20(3):687–706, Sept. 2017. ISSN 1573-6938. doi: 10.1007/s10683-016-9506-z. URL <https://doi.org/10.1007/s10683-016-9506-z>.
- S. L. Cheung, A. Tymula, and X. Wang. Present bias for monetary and dietary rewards. *Experimental Economics*, Mar. 2022. ISSN 1573-6938. doi: 10.1007/s10683-022-09749-8. URL <https://doi.org/10.1007/s10683-022-09749-8>.
- J. Cohen, K. M. Ericson, D. Laibson, and J. M. White. Measuring Time Preferences. *Journal of Economic Literature*, 58(2):299–347, June 2020. ISSN 0022-0515. doi: 10.1257/jel.20191074. URL <https://www.aeaweb.org/articles?id=10.1257/jel.20191074>.
- M. Collier and M. B. Williams. Eliciting individual discount rates. *Experimental Economics*, 2:107–127, 1999.
- R. P. Cubitt and D. Read. Can intertemporal choice experiments elicit time preferences for consumption? *Experimental Economics*, 10(4):369–389, Dec. 2007. ISSN 1573-6938. doi: 10.1007/s10683-006-9140-2. URL <https://doi.org/10.1007/s10683-006-9140-2>.
- D. J. Freeman. Revealing naïveté and sophistication from procrastination and pre-properation. *American Economic Journal: Microeconomics*, 13(2):402–438, 2021.

- F. Gul and W. Pesendorfer. Temptation and self-control. *Econometrica*, 69(6):1403–1435, 2001.
- Y. Halevy. Time consistency: Stationarity and time invariance. *Econometrica*, 83(1): 335–352, 2015.
- S. Kaur, M. Kremer, and S. Mullainathan. Self-control at work. *Journal of Political Economy*, 123(6):1227–1277, 2015.
- D. Laibson. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2):443–478, May 1997. ISSN 0033-5533. doi: 10.1162/003355397555253. URL <https://academic.oup.com/qje/article/112/2/443/1870925>.
- S. M. McClure, K. M. Ericson, D. I. Laibson, G. Loewenstein, and J. D. Cohen. Time discounting for primary rewards. *Journal of neuroscience*, 27(21):5796–5804, 2007.
- T. O’Donoghue and M. Rabin. Doing It Now or Later. *American Economic Review*, 89(1):103–124, Mar. 1999. ISSN 0002-8282. doi: 10.1257/aer.89.1.103. URL <https://www.aeaweb.org/articles?id=10.1257/aer.89.1.103>.
- T. O’Donoghue and M. Rabin. Choice and Procrastination. *The Quarterly Journal of Economics*, 116(1):121–160, 2001. ISSN 0033-5533. URL <https://www.jstor.org/stable/2696445>.
- E. S. Phelps and R. A. Pollak. On Second-Best National Saving and Game-Equilibrium Growth. *Review of Economic Studies*, 35(2):185–199, 1968. ISSN 0034-6527. doi: 10.2307/2296547. URL <https://www.jstor.org/stable/2296547>.
- D. Read and B. Van Leeuwen. Predicting hunger: The effects of appetite and delay on choice. *Organizational behavior and human decision processes*, 76(2):189–205, 1998.
- F. Schilbach. Alcohol and self-control: A field experiment in india. *American economic review*, 109(4):1290–1322, 2019.



- J. V. Solnick, C. H. Kannenberg, D. A. Eckerman, and M. B. Waller. An experimental analysis of impulsivity and impulse control in humans. *Learning and Motivation*, 11(1):61–77, 1980.
- S. Toussaert. Eliciting temptation and self-control through menu choices: A lab experiment. *Econometrica*, 86(3):859–889, 2018.

## 6 Appendix: Additional results

### 6.1 Structural estimates at the aggregate level, dropping the top 5% violators

Table 5: Structural estimates at the aggregate level

	[1] Same (AR2019)	[2] Separate	[3] No prediction	[4] Separate, $\delta = 1$
Present bias ( $\beta$ )	0.68 (0.51,0.85)			
Sophistication ( $\hat{\beta}$ )	0.73 (0.57,0.90)			
Present bias, effort ( $\beta_e$ )		0.67 (0.50,0.83)	0.67 (0.51,0.84)	0.79 (0.66,0.92)
Present bias, money ( $\beta_m$ )		0.96 (0.67,1.24)	0.94 (0.64,1.25)	1.18 (0.97,1.39)
Sophistication, effort ( $\hat{\beta}_e$ )		0.79 (0.55,1.02)		0.91 (0.68,1.14)
Sophistication, money ( $\hat{\beta}_m$ )		0.98 (0.72,1.25)		1.13 (0.88,1.37)
Discount factor $\delta$	1.27 (1.08,1.45)	1.15 (0.97,1.34)	1.15 (0.97,1.33)	
Convexity $\gamma$	2.50 (2.41,2.58)	2.57 (2.47,2.67)	2.57 (2.43,2.71)	2.56 (2.45,2.66)
Scale of effort $\psi$	199.78 (145.97,253.60)	207.61 (140.79,274.44)	211.71 (118.56,304.87)	199.90 (135.72,264.08)
Size of error $\sigma$	3.26 (3.20,3.32)	3.25 (3.19,3.31)	3.26 (3.19,3.33)	3.25 (3.19,3.31)
$\ln \mathcal{L}$	-6881.9	-6872.58	-4593.66	-6874.36
$N$	91	91	91	91
Observations	4095	4095	2730	4095

[1] *Same* follows [Augenblick and Rabin \(2019\)](#) to assume that present bias or sophistication over effort and money are identical. [2] *Separate* generalizes *Same* by allowing for different present bias and sophistication parameters across effort and money domains. [3] *No prediction* is similar to *Separate*, but excludes predictions and hence cannot estimate sophistication. [4] *Separate*,  $\delta = 1$  imposes that the long-run discount factor is one. The numbers in parentheses indicate the lower and upper bound of the 95% confidence interval.

### 6.2 Aggregate $\delta$

$e_{11}$  and  $e_{22}$  are decisions made about immediate effort, and the only difference is that for  $e_{22}$  the payment is a day closer than for  $e_{11}$ . Thus, irrespective of  $\beta_e, \beta_m, \hat{\beta}_e, \hat{\beta}_m$ ,

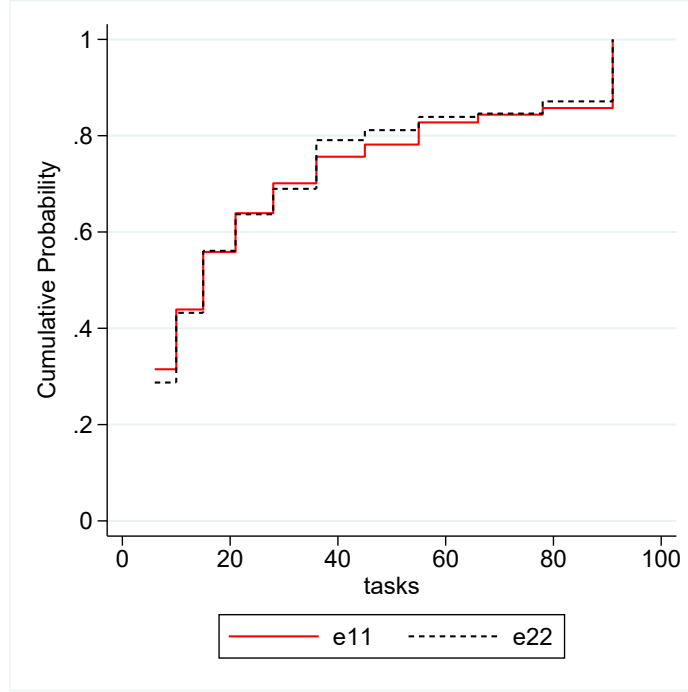


Figure 8: The empirical CDFs of  $e_{11}$  versus  $e_{22}$ .

under  $\delta < 1$  one would expect that  $e_{11} < e_{22}$ , and under  $\delta \geq 1$  one would expect that  $e_{11} \geq e_{22}$ . We plot the empirical CDFs in Figure 8. Even though the distributions are not statistically different under the Kolmogorov-Smirnov test,  $e_{11}$  lies to the right of  $e_{22}$  everywhere but at the lowest level of effort. This hints at  $\delta \geq 1$ , without being significantly different than 1.

### 6.3 Relative sophistication measure

Augenblick and Rabin (2019) quantify the relative sophistication of individuals by the quantity

$$\lambda = \frac{1 - \hat{\beta}_i}{1 - \beta_i} \quad (5)$$

and regress  $1 - \beta_i$  on  $1 - \hat{\beta}_i$  to calculate the average degree of understanding of present bias.<sup>24</sup> When we run the same OLS regression with all our 80 participants, we find that participants understand 26 percent of the present bias in the effort

<sup>24</sup>The regression coefficient measures the increase in  $1 - \hat{\beta}$  induced by a unit increase in  $1 - \beta$ .

domain ( $p < 0.000$ ).<sup>25</sup> The regression coefficient for sophistication over money is not significant.

## 7 Instructions

The next pages contain screenshots of the instructions as seen by our participants. The survey dynamically assigns the current date on the day of access, which is why all mentions of "current date/ today" in the survey show up as October 6th, 2023, the day when the survey was printed from Qualtrics.

---

<sup>25</sup>If we use robust standard errors, we get the same results.

**Week0 Introduction 1 with Consent****Welcome!**

Thank you for participating in this study.

Please read these instructions carefully. There will be a quiz later to make sure you understood the instructions.

If you have a question today, please write to us on Prolific. For later questions, please contact [c.cerrone@mdx.ac.uk](mailto:c.cerrone@mdx.ac.uk)

**Study details**

This study takes place on three days:

- today: Friday, October 6th (**Day 1**)
- tomorrow: Saturday, October 7th (**Day 2**) and
- the day after tomorrow: Sunday, October 8th (**Day 3**)

**Please participate only if you can attend on all three study days!**

On all three days,

- You will need to complete the study between 7 am and 7 pm.
- You must use a computer (not a tablet or mobile phone).

Participation today will take at least 15 minutes (about 10 minutes for the instructions and 5 minutes for the experiment), but it can take up to 45 minutes (in that case you would also earn more money).

Participation tomorrow and on the day after tomorrow will take at least 5 minutes. It can take longer if you decide to work for additional money.

We will use the data from this experiment for academic research only. Your anonymity is guaranteed and we will not ask you to provide personal data.

### Consent

I hereby agree to participate in this study on all 3 days. My anonymized data can be used for scientific publications.

- ☐ I agree
- ☐ I do not agree

### Does not consent

***“As you do not wish to participate in this study, please return your submission on Prolific by selecting the 'Stop without completing' button.”***

### Week0 Introduction after Consent

#### Your Prolific ID

What is the Prolific ID you used to log into this study?

### Payment details

In this study, you can get paid for 4 things.

**1. Baseline payments.** The baseline payment for today is £2 and the baseline payments for tomorrow and the day after tomorrow are £1 each.

**2. Bonus payment of £5 for completing the study.** If you successfully participate in all 3 study days, you will receive £5 as a completion payment.

**3. Bonus payments for completing "transcription tasks".** On each of the 3 study days you have the opportunity to earn additional money by completing "transcription tasks". We will explain these tasks on the next page.

**4. Additional bonus payments.** We will explain this later.

**Important:** If you miss one of the 3 study days, you will not receive the £5 completion payment and you will not participate in the following study days. You will only be paid for the "transcription tasks" that you have already completed. We cannot allow exceptions to this rule.

We will pay you your total earnings from the study on **the day after tomorrow** (Sunday, October 8th), after you complete the last survey. All payments will be made on that day.

### Choosing Transcription Tasks and Jobs

In this study, we will ask you how many "transcription tasks" and "jobs" you want to complete for different wages at different times.

**Relationship between Transcription Tasks and Jobs** The number of transcription tasks you need to complete for each number of jobs is calculated in the following table.

Jobs Chosen	Conversion	Transcription Tasks
1	1	1
2	1+2	3
3	1+2+3	6
4	1+2+3+4	10

5	1+2+3+4+5	15
6	1+2+3+4+5+6	21
7	1+2+3+4+5+6+7	28
8	1+2+3+4+5+6+7+8	36
9	1+2+3+4+5+6+7+8+9	45
10	1+2+3+4+5+6+7+8+9+10	55
11	1+2+3+4+5+6+7+8+9+10+11	66
12	1+2+3+4+5+6+7+8+9+10+11+12	78
13	1+2+3+4+5+6+7+8+9+10+11+12+13	91

Whenever you choose the number of jobs to complete in the following pages, we will show you the corresponding number of transcription tasks. You do not have to calculate it by yourself.

**What is a Transcription Task?** Here is an example:

Please transcribe the symbols on the left into the blank field on the right.

Yve&S{xG=Vk{=t

If you make a mistake in the transcription, you will receive feedback and can correct the mistake immediately.

On average, participants complete about 3 of these tasks in a minute.

### Decisions about transcription tasks

You decide how many of these jobs (from 3 to 13) you want to complete. You can choose to complete a different number of jobs on different days and at different wages. You have to do at least 3 jobs (6 transcription tasks) and 13 jobs (91 transcriptions tasks) at the maximum.



**Your decisions determine how much you will work and get paid.** For each day, the survey software will randomly select ONE of your decisions. For this randomly selected decision, you will be asked to complete your chosen number of tasks at the relevant wage. As a reminder, the number of transcription tasks you need to complete will be adding all numbers between 1 and the number of jobs you have chosen.

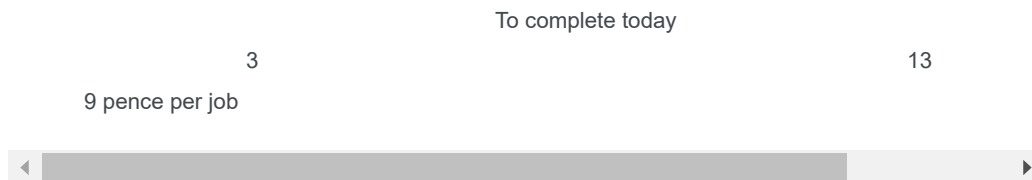
If the number of jobs chosen is ***J***, and the wage rate chosen is ***W***, your total wage will be ***J* times *W*** for that day

Below is an example with a wage of 9 pence per job. In the real experiment, you will see several of these sliders.

### Jobs to be completed **TODAY (Example Question)**

How many transcription jobs do you want to complete **today (in this session)** for the given wage? Clicking the slider also shows the number of transcription tasks you need to complete.

Note: Please click the slider again if the numbers are not shown properly.



As you can see, you can select the number of jobs by moving the sliders between 3 and 13.

You will face wages between 1 penny per job and 9 pence per job. You only get the completion payment if you correctly type in at least 95% of the tasks you committed to complete.

### Week0 Quiz

## Quiz

Please answer all the questions in this quiz.

On how many days do you have to participate in this study, including today?

- ☐ Only today
- ☐ 2 days
- ☐ 3 days
- ☐ 4 days
- ☐ I can choose

What is the completion payment for participating in all 3 days of the study?

- ☐ £1
- ☐ £3
- ☐ £5
- ☐ £10

When will you receive payments for participating in this study?

- ☐ Today (on Day 1)
- ☐ Tomorrow (on Day 2)
- ☐ The day after tomorrow (on Day 3)
- ☐ On every study day

## Training the tasks

Yve&amp;S{xG=Vk{=t

z9zK9]8"R1MzCP

ZgSXbR[QM3zchQ

## Predictions

### Predict your choices tomorrow

The questions below and on the next page ask you to make predictions about how you will decide tomorrow.

The survey software will randomly choose one of these predictions and compare it with your actual decision. You will earn 20 pence if the randomly selected prediction task is within 2 jobs of your actual decision.

### Prediction 1

How many transcription jobs do you think you will choose **tomorrow** to complete **tomorrow (in the same session)** (for the 5 different wages)?

Note: Please click the slider again if the numbers are not shown properly.

Predicted decision tomorrow for tasks to complete on the same day

3

13

1 penny per job

3 pence per job

Predicted decision tomorrow for tasks to complete on the same day

3

13

5 pence per job

7 pence per job

9 pence per job

## Prediction 2

How many transcription jobs do you think you will choose **tomorrow** to complete on **the day after tomorrow** (for the 5 different wages)?

Note: Please click the slider again if the numbers are not shown properly.

Predicted decision tomorrow for tasks to complete on the day after tomorrow

3

13

1 penny per job

3 pence per job

5 pence per job

7 pence per job

9 pence per job



## Week0 Demographics

What is your gender?

☐ Male

☐ Female

- ☐ Other (e.g. not binary; specification is voluntary)
- ☐ Skip question

What is your marital status?

- ☐ Single
- ☐ Married / living together
- ☐ Divorced
- ☐ Other (please specify )

If you had to, how much money could you spend today without having to limit yourself financially in the next 4 weeks.

- ☐ £0
- ☐ £0.01 - £20
- ☐ £20.01 - £50
- ☐ £50.01 - £100
- ☐ £100.01 - £200
- ☐ £200.01 - £500
- ☐ £500.01 - £1000
- ☐ £1000.01 and above

**Week 1 Thanks**

**Thank you!**

This concludes today's part of the study.

***Note: You must submit this survey by clicking on the right arrow below for your data to be transferred to tomorrow's survey.***

This is a pilot study. Do you have any feedback for us?

## **Week 2 Welcome**

### **Welcome back to the 2nd day of the study!**

Today (Friday, October 6th) is the second day of the study. The study will end tomorrow (Saturday, October 7th).

Here is an overview of today's study session:

1. Choose tasks to be completed today and tomorrow.
2. Perform the tasks for this week.
3. Predict the decisions that you will make tomorrow.

Recall that:

- You must use a computer (not a tablet or mobile phone).
- Participation can take up to 45 minutes (if you decide to complete many tasks).
- We will use the data from this experiment for academic research only. Your anonymity is guaranteed and we will not ask you to provide personal data.

### **Consent**

I hereby agree to participate in this study today and tomorrow again. My anonymized data can be used for scientific publications.

- ☐ I agree
- ☐ I do not agree

### **Do not consent**

***“As you do not wish to participate in this study, please return your submission on Prolific by selecting the 'Stop without completing' button.”***

### **Prolific and payment details**

#### **Your Prolific ID**

What is the Prolific ID you used to log into this study?

(We already used your Prolific ID to transfer the data from yesterday to today's survey.)

#### **Reminder of the payment details**

In this study, you can get paid for 4 things.

- 1. Baseline payments.** The baseline payment for yesterday was £2 and the baseline payments for today and tomorrow are £1 each.
- 2. Bonus payment of £5 for completing the study.** If you successfully participate in all 3 study days, you will receive £5 as a completion payment.
- 3. Bonus payments for completing "transcription tasks".** On each of the 3 study days you have the opportunity to earn additional money by completing "transcription tasks".
- 4. Bonus payments for correct predictions.** You will make predictions about your own future decisions. One of these predictions will be randomly chosen. If the prediction is accurate, you will get 20 pennies.

**Important:** If you miss one of the 3 study days, you will not receive the £5 completion payment and you will not be able to participate in the following days. You will only be paid

for the "transcription tasks" that you have already completed. We cannot allow exceptions to this rule.

We will pay you tomorrow (Saturday, October 7th). All payments will be made on that day.

## Week 2 Decisions and Predictions

### Decisions about transcription tasks

On this and the next page, you can decide how many transcription tasks you want to complete for different wages.

Make each decision carefully. The survey software will randomly choose one of your decisions and ask you to complete your chosen number of tasks at the relevant wage.

### Jobs to complete **TOMORROW ON Saturday, October 7th**

How many transcription jobs do you want to complete **tomorrow (Saturday, October 7th)** for the 5 different wages?

Reminder: Only ONE of your decisions will be selected and implemented. This determines how many tasks you must complete tomorrow.

Note: Please click the slider again if the numbers are not shown properly.

To complete tomorrow

3 13

1 penny per job

3 pence per job

5 pence per job



To complete tomorrow

3

13

7 pence per job

9 pence per task

**Jobs to complete TODAY IN A FEW MOMENTS**

How many transcription jobs do you want to complete **today (in this session)** for the 5 different wages?

Reminder: Only ONE of your decisions will be selected and implemented. This will determine how many tasks you have to complete in a few moments.

Note: Please click the slider again if the numbers are not shown properly.

Today

3

13

1 penny per job

3 pence per job

5 pence per job

7 pence per job

9 pence per job

**Week 2 Pick the Task**

Below are the 5 decisions (1-5) you made yesterday about the number of jobs to complete today, and the 5 decisions (6-10) that you made today for the tasks to be completed today:

Decision	Decision Day	Pay per job	The number of jobs chosen	The number of tasks you want to perform
1	Yesterday	1 Penny	#{e://Field/j12_1}	#{e://Field/e12_1}
2	Yesterday	3 Pence	#{e://Field/j12_2}	#{e://Field/e12_2}
3	Yesterday	5 Pence	#{e://Field/j12_3}	#{e://Field/e12_3}
4	Yesterday	7 Pence	#{e://Field/j12_4}	#{e://Field/e12_4}
5	Yesterday	9 Pence	#{e://Field/j12_5}	#{e://Field/e12_5}
6	Today	1 Penny	#{e://Field/j22_1}	#{e://Field/e22_1}
7	Today	3 Pence	#{e://Field/j22_2}	#{e://Field/e22_2}
8	Today	5 Pence	#{e://Field/j22_3}	#{e://Field/e22_3}
9	Today	7 Pence	#{e://Field/j22_4}	#{e://Field/e22_4}
10	Today	9 Pence	#{e://Field/j22_5}	#{e://Field/e22_5}

On the next page, we will announce the randomly selected decision that will determine how many tasks you must complete today. This decision was chosen at random for you by the survey software.

The randomly selected decision for today is "\${e://Field/Random1-14Wednesday}".

Decision	Decision Day	Pay per task	The number of jobs chosen	The number of tasks you want to perform
1	Yesterday	1 Penny	\${e://Field/j12_1}	\${e://Field/e12_1}
2	Yesterday	3 Pence	\${e://Field/j12_2}	\${e://Field/e12_2}
3	Yesterday	5 Pence	\${e://Field/j12_3}	\${e://Field/e12_3}
4	Yesterday	7 Pence	\${e://Field/j12_4}	\${e://Field/e12_4}
5	Yesterday	9 Pence	\${e://Field/j12_5}	\${e://Field/e12_5}
6	Today	1 Penny	\${e://Field/j22_1}	\${e://Field/e22_1}
7	Today	3 Pence	\${e://Field/j22_2}	\${e://Field/e22_2}
8	Today	5 Pence	\${e://Field/j22_3}	\${e://Field/e22_3}
9	Today	7 Pence	\${e://Field/j22_4}	\${e://Field/e22_4}
10	Today	9 Pence	\${e://Field/j22_5}	\${e://Field/e22_5}

This means that you have \${e://Field/TasksDay2} transcription tasks to complete today, and your wage per each job is \${e://Field/WageDay2} penny/pence.

These \${e://Field/TasksDay2} tasks are shown on the next page.

Av?s)iJQR9Cjef

Yve&amp;S{xG=Vk{=t

z9zK9]8"R1MzCP

ZgSXbR[QM3zchQ

### Prediction Task

Predict your choices **TOMORROW**

One of the predictions you make in this study will be randomly chosen. If this prediction is accurate, you will get 20 pence.

### Prediction

How many transcription jobs do you think you **will** choose **TOMORROW, Saturday, October 7th, to complete in the same session** (for the 5 different wages)?

Note: Please click the slider again if the numbers are not shown properly.

Prediction: decision for tasks tomorrow to complete in the same session

3

13

1 penny per job

3 pence per job      Prediction: decision for tasks tomorrow to complete in the same session

5 pence per job 3

13

7 pence per job

9 pence per job



## Week 2 Thanks

**Thank you!**

This concludes today's part of the study.

***Note: You must submit this survey by clicking on the right arrow below for your data to be transferred to next week's survey.***

This is a pilot study. Do you have any feedback for us?

Just click below to go back to Prolific.

Powered by Qualtrics

### **Week 3 Mandatory Survey**

#### **Welcome back to the last day of the study!**

Today (Friday, October 6th) is the third day of the study. The study will end today and you will be paid today as well.

Here is an overview of today's study session:

1. Choose tasks to be completed today
2. Perform the tasks for today.
3. Provide some feedback on how the survey could be improved.

Recall that:

- You must use a computer (not a tablet or mobile phone).
- Participation can take up to 45 minutes (if you decide to complete many tasks).
- We will use the data from this experiment for academic research only. Your anonymity is guaranteed and we will not ask you to provide personal data.

#### **Consent**

I hereby agree to participate in this study today. My anonymized data can be used for scientific publications.

- ☐ I agree
- ☐ I do not agree

#### **Does not consent**

***“As you do not wish to participate in this study, please return your submission on Prolific by selecting the 'Stop without completing' button.”***

### Intro after Consent

#### Your Prolific ID

What is the Prolific ID you used to log into this study?

(We already used your Prolific ID to transfer the data from yesterday to today's survey.)

#### Reminder of the payment details

In this study, you can get paid for 4 things.

**1. Baseline payments.** The baseline payment for the first day and yesterday were £2 and £1 each. The baseline payments for today is £1.

**2. Bonus payment of £5 for completing the study.** If you successfully participate in all 3 study days, you will receive £5 as a completion payment.

**3. Bonus payments for completing "transcription tasks".** On each of the 3 study days you have the opportunity to earn additional money by completing "transcription tasks".

**4. Bonus payments for correct predictions.** You have made predictions about your own future decisions. One of these predictions will be randomly chosen. If the prediction is accurate, you will get 20 pence.

**Important:** If you miss one of the 3 study days, you will not receive the £5 completion payment and you will not be able to participate in the following days. You will only be paid

for the "transcription tasks" that you have already completed. We cannot allow exceptions to this rule.

All payments will be made today.

### Week 3 Decisions

#### Decisions about transcription tasks

On this page, you can decide how many transcription tasks you want to complete for different wages.

Make each decision carefully. The survey software will randomly choose one of your decisions and ask you to complete your chosen number of tasks at the relevant wage.

#### Jobs to complete **TODAY IN A FEW MOMENTS**

How many transcription jobs do you want to complete **today (in this session)** for the 5 different wages?

Reminder: Only ONE of your decisions will be selected and implemented. This determines how many tasks you must complete in a few moments.

Note: Please click the slider again if the numbers are not shown properly.

	To complete in a few moments	
	3	13
1 penny per job		
3 pence per job		
5 pence per job		
7 pence per job		



To complete in a few moments

3

13

9 pence per job

**Week 3 Pick the Task**

Below are the 15 decisions you made for today (on the day before yesterday, yesterday, and today):

Decision	Decision Day	Pay per job	The number of jobs chosen	The number of tasks you want to perform
1	Day before yesterday	1 Penny	$\{e://Field/j13\_1\}$	$\{e://Field/e13\_1\}$
2	Day before yesterday	3 Pence	$\{e://Field/j13\_2\}$	$\{e://Field/e13\_2\}$
3	Day before yesterday	5 Pence	$\{e://Field/j13\_3\}$	$\{e://Field/e13\_3\}$
4	Day before yesterday	7 Pence	$\{e://Field/j13\_4\}$	$\{e://Field/e13\_4\}$
5	Day before yesterday	9 Pence	$\{e://Field/j13\_5\}$	$\{e://Field/e13\_5\}$
6	Yesterday	1 Penny	$\{e://Field/j23\_1\}$	$\{e://Field/e23\_1\}$

7	Yesterday	3 Pence	$\{e://Field/j23\_2\}$	$\{e://Field/e23\_2\}$
8	Yesterday	5 Pence	$\{e://Field/j23\_3\}$	$\{e://Field/e23\_3\}$
9	Yesterday	7 Pence	$\{e://Field/j23\_4\}$	$\{e://Field/e23\_4\}$
10	Yesterday	9 Pence	$\{e://Field/j23\_5\}$	$\{e://Field/e23\_5\}$
11	Today	1 Penny	$\{e://Field/j33\_1\}$	$\{e://Field/e33\_1\}$
12	Today	3 Pence	$\{e://Field/j33\_2\}$	$\{e://Field/e33\_2\}$
13	Today	5 Pence	$\{e://Field/j33\_3\}$	$\{e://Field/e33\_3\}$
14	Today	7 Pence	$\{e://Field/j33\_4\}$	$\{e://Field/e33\_4\}$
15	Today	9 Pence	$\{e://Field/j33\_5\}$	$\{e://Field/e33\_5\}$

On the next page, we will announce the randomly selected decision that will determine how many tasks you must complete today. This decision was chosen at random for you by the survey software.

**The randomly selected decision for today is " $\{e://Field/Random1-21Thursday\}$ ".**

Decision	Decision Day	Pay per job	The number of jobs chosen	The number of tasks you want to perform
1	Day before yesterday	1 Penny	$\{e://Field/j13\_1\}$	$\{e://Field/e13\_1\}$
2	Day before yesterday	3 Pence	$\{e://Field/j13\_2\}$	$\{e://Field/e13\_2\}$
3	Day before yesterday	5 Pence	$\{e://Field/j13\_3\}$	$\{e://Field/e13\_3\}$
4	Day before yesterday	7 Pence	$\{e://Field/j13\_4\}$	$\{e://Field/e13\_4\}$
5	Day before yesterday	9 Pence	$\{e://Field/j13\_5\}$	$\{e://Field/e13\_5\}$
6	Yesterday	1 Penny	$\{e://Field/j23\_1\}$	$\{e://Field/e23\_1\}$
7	Yesterday	3 Pence	$\{e://Field/j23\_2\}$	$\{e://Field/e23\_2\}$
8	Yesterday	5 Pence	$\{e://Field/j23\_3\}$	$\{e://Field/e23\_3\}$
9	Yesterday	7 Pence	$\{e://Field/j23\_4\}$	$\{e://Field/e23\_4\}$
10	Yesterday	9 Pence	$\{e://Field/j23\_5\}$	$\{e://Field/e23\_5\}$
11	Today	1 Penny	$\{e://Field/j33\_1\}$	$\{e://Field/e33\_1\}$
12	Today	3 Pence	$\{e://Field/j33\_2\}$	$\{e://Field/e33\_2\}$
13	Today	5 Pence	$\{e://Field/j33\_3\}$	$\{e://Field/e33\_3\}$
14	Today	7 Pence	$\{e://Field/j33\_4\}$	$\{e://Field/e33\_4\}$
15	Today	9 Pence	$\{e://Field/j33\_5\}$	$\{e://Field/e33\_5\}$

This means that you have to do  $\${e://Field/TasksToDoED}$  transcription tasks. Your pay for each job is  $\${e://Field/WageDay3}$  penny/pence.

These  $\${e://Field/TasksToDoED}$  tasks are shown on the next page.

### Perform 100 Tasks

#### Day 3 Task

You don't have any tasks to do today.

riCs&XGq5rehf8

hQYVMLLJ6!kAq8

J3uyYwrC7B!Pt1

P)KzUB8aj9b1??

\6zHYFQew}Ry1h

Av?s)iJQR9Cjef

Yve&S{xG=Vk{=t

z9zK9]8"R1MzCP

ZgSXbR[QM3zchQ

### Week 3 Thanks

### Thank you!

This concludes the study. We will make all payments today. We would appreciate very much if you could answer the questions below.

How did you predict your future choices on days 1 and 2?

Did you find the instructions clear? Please tell us how it could be improved.

Did you use any strategy to ensure that your predictions better matched your own future choices? If so, what strategy did you use?

Is there anything else you would like to share with us?

Just click below to go back to Prolific.

Powered by Qualtrics