

# Why Do We Procrastinate? Present Bias and Optimism

Zachary Breig, Matthew Gibson, and Jeffrey G. Shrader\*

April 23, 2024

## Abstract

Time inconsistency is typically modeled as resulting from present bias, a form of preference non-stationarity. In this paper, we study an alternative: excessively optimistic beliefs about future costs or demands on an individual's time. The models can be distinguished by how individuals respond to information. Experimental results refute the hypothesis that non-stationarity is the sole source of dynamic inconsistency, but they are consistent with biased beliefs about shocks. These findings offer an explanation for low takeup of commitment and suggest that personalized information can mitigate procrastination.

JEL: C91, D15, D84, D90, J22

Keywords: dynamic inconsistency, beliefs, discounting, preference non-stationarity

---

\*Affiliations. Breig: School of Economics, The University of Queensland, Level 6 Colin Clark Building (39), St Lucia, Brisbane, Australia, 4072 (e-mail: [z.breig@uq.edu.au](mailto:z.breig@uq.edu.au)); Gibson: Department of Economics, Williams College, 24 Hopkins Hall Drive, Williamstown MA 01267, and IZA (e-mail: [mg17@williams.edu](mailto:mg17@williams.edu)); Shrader: School of International and Public Affairs, Columbia University, 420 West 118th Street, New York NY 10027 (e-mail: [jgs2103@columbia.edu](mailto:jgs2103@columbia.edu)). We are grateful for helpful comments from James Andreoni, Gordon Dahl, Benjamin Enke, Marco Faravelli, Osea Giuntella, Jeff Kline, Nicholas Lacourse, Paul Lohmann, Bentley MacLeod, Frank Schilbach, Charles Sprenger, and Jack Willis as well as seminar participants at the ASSA annual meeting, ASHEcon, the North American ESA Meeting, UCSD, the University of Melbourne, and the University of Queensland. George Kassabian and Anmolbir Mann provided excellent research assistance. Funding was provided by Columbia University and Williams College. This study was approved by the IRBs at the University of California, San Diego, Columbia University, and the University of Queensland. The Online Experiment described in this paper was registered as AEARCTR-0011140 (Breig et al., 2023).

## 1 Introduction

Procrastination is an important feature of everyday life. It is a common topic of conversation at work and at home, and economists have documented it in consequential settings including retirement saving, exercise, and education (DellaVigna, 2009). Procrastination is commonly modeled as originating from present biased discounting that favors the present at the expense of the future (Strotz, 1955, Laibson, 1997, O’Donoghue and Rabin, 1999, Barro, 1999, Ashraf et al., 2006, Heidhues and Köszegi, 2010, Augenblick et al., 2015).<sup>1</sup> We study an alternative model in which dynamic inconsistency arises from excessive optimism about future demands on an individual’s time. While both models predict dynamically inconsistent choices, they predict different responses to information that changes beliefs. We test these predictions experimentally and reject the hypothesis that present bias is the sole source of dynamic inconsistency. Instead we find evidence that biased beliefs about demands on time matter. Our results suggest that the typical policy prescription—committing to decisions in advance—is incomplete and that personalized historical information is an important additional tool for people making decisions over time.

Biased beliefs about one’s cost or time shocks can cause choices made ahead of time to differ from choices made in the moment. Consider an agent who does not accurately anticipate the arrival of a time-consuming task. Colloquially we say that such an agent is optimistic about her time shocks. Once the task arrives, the agent will need to reallocate her planned time use to accommodate the unanticipated shock.<sup>2</sup> If the agent has systematically biased beliefs over future time shocks, then such procrastination can occur even with neoclassical discounting. We refer to this source of dynamic inconsistency as *biased beliefs about shocks*.<sup>3</sup>

Dynamic inconsistency can also arise from preference non-stationarity, which does not require that preferences between options remain unchanged if the timing of all options is shifted by the same amount (Koopmans, 1960, Halevy, 2015). This leads the

---

<sup>1</sup>In the quasi-hyperbolic model of Laibson (1997), the agent discounts at rate  $\delta$  between future periods, but between the current period and the next period at rate  $\beta\delta$  with  $\beta < 1$ . This heavier discounting leads to “present biased” allocative choices. Chakraborty (2021) axiomatizes a model of present bias and shows that a variety of models of time preferences satisfy the axioms.

<sup>2</sup>The model shares features with Kahneman and Tversky’s (1982) “planning fallacy.” Our theory links biased beliefs and dynamic inconsistency.

<sup>3</sup>In contrast to Halevy (2008) and Andreoni and Sprenger (2012), this inconsistency is a result of the decision maker having incorrect beliefs rather than a utility function that does not take the expected utility form.

agent to exhibit dynamic inconsistency because choices made far enough in advance will be governed by one set of preferences, while choices made about the immediate future will be governed by a different set. If an agent is naïve about her own non-stationarity, she believes that she will behave more consistently than she actually does, so we refer to this source of dynamic inconsistency as *naïve non-stationarity*.<sup>4</sup>

Because these two models lead to similar dynamically inconsistent choices, a research design seeking to distinguish naïve non-stationarity from biased beliefs about shocks cannot rely solely on revealed procrastination. However, with assumptions about how subjects update their beliefs, making past behavior salient to decision makers resolves this identification problem. Specifically, we assume that when confronted with past procrastination, a decision maker with biased beliefs about shocks updates those beliefs to be more pessimistic. In the same situation, we assume that a decision maker with naïve nonstationarity would become “more sophisticated,” increasing the likelihood of procrastination according to the decision maker’s beliefs. Thus, both models rely on agents not incorporating information about their prior choices when forming beliefs, but the two models make different predictions about how agents will respond to information that changes beliefs.

First, the two models give different predictions for how effort allocation will change in response to reminders about past dynamic inconsistency. Naïve non-stationary agents have a clear idea of the shocks that they face, but have trouble committing to time use choices. Such agents will not change effort allocations in response to information. In contrast, for agents with biased beliefs about shocks, correcting these beliefs will cause them to change their effort allocations to better conform to the true state of the world.

Second, information can cause naïve non-stationary agents to learn about their own non-stationarity. For instance, an agent might learn that her discounting is more present-biased than she previously thought. This will increase commitment demand for time-use choices made far enough in advance. If agents have biased beliefs over time shocks, however, this prediction need not hold. Information on past dynamically inconsistent decisions should help belief-based dynamically inconsistent agents bring their beliefs in line with the true state, but this does not necessarily lead them to demand costly commitment (Laibson, 2015).<sup>5</sup>

---

<sup>4</sup>In the quasi-hyperbolic model, partially naïve agents have true discounting parameters  $\beta$  and  $\delta$  but believe their present-bias parameter is  $\hat{\beta}$  where  $\beta < \hat{\beta} \leq 1$  (O’Donoghue and Rabin, 2001).

<sup>5</sup>Optimistic agents with an underlying neoclassical utility function would generally like flexibility

We tested these predictions in two experiments. Both lasted for two weeks, with the first week allowing us to measure baseline dynamically inconsistent behavior for each subject. In both weeks, subjects chose how many real-effort tasks to complete at a fixed date. In the first experiment, each task done at the beginning of the week reduced the number of tasks that needed to be completed later in the week, while in the second experiment, subjects were paid on a future date for each task completed.<sup>6</sup> Subjects made each choice twice: once well in advance of task completion, and once immediately before. Subjects could reveal a preference for costly commitment by choosing whether the earlier or later choice would be more likely to be implemented using a price list denominated in additional tasks.

At the beginning of week 2, before engaging in task decisions like those in week 1, treated subjects were presented with information. In the first experiment, they were presented with a reminder about their task choices in the first week.<sup>7</sup> In the second experiment, one set of subjects was treated with reminders about their task choices in the first week, while the other set of subjects was given information relevant to how difficult the tasks would be in week 2. Because we only treated some subjects with this information, the experiments allow us to identify the effect of changing beliefs while controlling for other determinants of procrastination and commitment demand.

Our experimental results indicate that both biased beliefs about shocks and non-stationary preferences are important determinants of time inconsistency. The strongest evidence is on biased beliefs about shocks and comes from testing the effect of treatment on task allocation. In both experiments we observed reallocation, on average, in response to each of the treatments. Among subjects who reallocated work in week 1, the treatment with information caused a practically large and statistically significant change in reallocation of tasks in week 2. This is inconsistent with preference non-stationarity and consistent with biased beliefs over cost shocks. Evidence on naive non-stationarity comes from testing the effect of treatment on commitment demand in week 2 for individuals who reallocated in week 1. In the first experiment, treatment increased week 2 commitment demand among reallocators, while in the second experiment it did not.<sup>8</sup>

---

and making beliefs less optimistic has an ambiguous effect on commitment demand.

<sup>6</sup>A discussion of the motivation for the differences between the two experiments is in Section 2.3.4.

<sup>7</sup>In the first experiment, subjects were also given information about how well they were able to forecast their own bedtimes, a real-world procrastination behavior.

<sup>8</sup>Section 3.2.3 discusses potential explanations for this difference in results.

Our first experiment shares design features with [Augenblick et al. \(2015\)](#), which is part of an extensive literature evaluating the prevalence of dynamic inconsistency across a variety of domains. [Augenblick et al. \(2015\)](#) shows that in their experiment, present bias is more common in real-effort choices than in choices over time-dated money. It also demonstrates that observed present bias is correlated with the take up of commitment. In contrast, our study is focused on identifying the sources of dynamic inconsistency rather than measuring it. Because of this, our paper is more closely related to [Halevy \(2015\)](#), which identifies how non-stationarity and time-varying preferences contribute to time inconsistency using a classroom-based experiment with time-dated monetary rewards. We add to this literature by showing that biased beliefs contribute to time inconsistency and that targeted information can reduce that time inconsistency.

This study provides empirical evidence that a model of dynamically inconsistent behavior based purely on naive non-stationarity is incomplete. The presence of biased beliefs about shocks matters for policy aimed at dynamically inconsistent behavior. The prescription from the time inconsistency literature has primarily been to encourage commitment by sophisticated present-biased agents. Our results suggest that such tools are inappropriate for some people. If procrastination stems from overestimation of future earnings or underestimation of how difficult it will be to quit smoking, then organizations and individuals seeking to correct dynamic inconsistency should provide personalized, salient information. This hypothesis is consistent with the widespread sale of goods—like fitness trackers and planners—that help consumers reflect on execution of their own plans.<sup>9</sup>

In addition, our study makes two contributions to research on demand for costly commitment. First, our findings help explain the widely observed low take-up of such commitment. Subjects whose dynamic inconsistency originates solely from optimism will not demand costly commitment. [Schilbach \(2019\)](#) observes that in the majority of past experiments, subjects were either unwilling to pay for commitment or were willing to pay only very small amounts. Second, our experimental design makes a methodological contribution in its elicitation of commitment demand. In contrast to most previous work eliciting commitment demand, our commitment price is denominated in tasks rather than money.<sup>10</sup> By keeping all choices in the task domain, we

---

<sup>9</sup>Paul Krugman has made this point when reflecting on his own fitness tracker use, writing that “what fitness devices do, at least for me, is make it harder to lie to myself” ([Krugman, 2015](#)).

<sup>10</sup>To the best of our knowledge [Toussaert \(2018\)](#) is the only other experiment that elicits commit-

reduce the tendency of commitment demand to spike sharply at a zero price. We find that about one-quarter of the subjects in the lab experiment and over one-half of the subjects in the online experiment were willing to commit to their time use choices at positive task-denominated prices.

Finally, our results contribute to a growing body of research demonstrating the importance of a decision maker’s beliefs for how they make choices involving time. There is evidence that decision makers are subject to the “planning fallacy” when forming beliefs about future events (Kahneman and Tversky, 1982, Roy et al., 2005). DellaVigna and Malmendier (2006) and Acland and Levy (2015) both study gym membership and attendance, showing that consumers systematically overestimate how often they will go to the gym in the future even when this choice entails monetary costs. Börsch-Supan et al. (2018) demonstrate that a much larger portion of regret about not having saved more earlier in life is explained by positive and negative financial shocks than present bias. Allcott et al. (2021) show that some payday loan borrowers exhibit evidence of both time inconsistency and overoptimism about repayment. Consistent with the common lack of commitment demand in experimental subjects, Augenblick and Rabin (2018) find that individuals’ *predictions* about the choices they will make in the future suggest that they do not understand their own present bias. Furthermore, subjects who make choices for the future immediately after completing tasks volunteer for less work in the future than those asked just before completing tasks. While the authors interpret this as evidence of projection bias, it is also consistent with decision makers who are optimistic about their desire to complete future tasks but who update after getting information. Our paper experimentally tests the link between time inconsistency and a planning fallacy that occurs due to biased beliefs about shocks.

The paper proceeds as follows. Section 2 describes the motivating theory, lays out testable hypotheses, and gives the experimental design. Section 3 presents empirical tests of our hypotheses. Section 4 concludes.

## 2 Experimental Design

In this section, we begin by describing our conceptual framework and providing experimental hypotheses. We then describe two longitudinal experiments that we use to study the sources of procrastination.

---

ment demand with prices denominated in tasks.

The first experiment (henceforth the “lab experiment”) was run with a student sample at a large university in the United States. The first session of the experiment, in which experimental procedures were explained, was run in person in a laboratory. All follow-up experimental responses were completed online using the survey software Qualtrics. As discussed in Section 2.3.4, some of the surprising results of the lab experiment motivated us to conduct a follow-up experiment with a larger sample, an additional treatment, and a slightly altered design.

The second experiment (henceforth the “online experiment”) was conducted with an online sample using the recruitment website Prolific.<sup>11</sup> After volunteering for the study on Prolific, subjects completed all experimental responses on Qualtrics.

## 2.1 Conceptual Framework and Hypotheses

Consider a decision maker who must choose how many tasks to complete at some date  $t$ . Completion of these tasks has some benefit at a later date, with more tasks completed bringing higher benefits. However, tasks are costly in terms of time and effort, and costs may be unknown before date  $t$ . We study two choices the decision maker might make: how she chooses well in advance of the tasks needing to be completed (henceforth “committed decisions”), and immediately before they are completed (“uncommitted decisions”). This paper focuses on why these choices might differ systematically. We call any difference in the chosen number of tasks a *reallocation*. Positive values for reallocation mean the number of tasks that the decision maker chose well in advance of completion is higher than the number she chose just before completion; negative values mean the opposite.

There are many reasons why committed decisions might be different from uncommitted decisions. The economics literature has placed significant emphasis on procrastination (a positive average value of reallocation). The most common explanation for these reallocations is non-stationary preferences, the leading model of which is  $\beta$ - $\delta$  (Laibson, 1997). In this model, the difference between committed and uncommitted decisions arises as a result of differences in the way the decision maker discounts. When the committed decision is made, both costs and benefits are in the future. When the uncommitted decision is made, costs are in the present and benefits are in the future, leading to benefits being systematically down-weighted and fewer

---

<sup>11</sup>Comparisons of responses by participants recruited from Prolific, other online platforms, and undergraduate subject pools show that Prolific subjects are typically more diverse and their participation is of higher quality (Eyal et al., 2021).



tasks being completed. Thus, choices exhibit present-bias. Furthermore, research has shown that to explain behavior, decision makers must be at least partially naive about their present bias: they act as if they believe they will behave in a less present-biased way than they do (Acland and Levy, 2015, Augenblick et al., 2015, Augenblick and Rabin, 2018, Le Yaouanq and Schwardmann, 2022).

In this paper, we propose an alternative explanation for systematic task reallocations: biased beliefs about costs.<sup>12</sup> When the decision maker makes the committed choice, she might not know exactly the costs that she will face when she completes the tasks. When she makes the uncommitted choice, she would likely be aware of any time shocks that are relevant to her costs. Biased beliefs about these shocks could then lead to systematically reallocating tasks.<sup>13</sup> For instance, if the decision maker is optimistic about the costs she will face, then she would tend to be surprised by high realizations of costs, leading to systematic task reallocation.

Because both naive non-stationarity and biased beliefs about shocks lead to systematic task reallocation, observing these task reallocations cannot differentiate the models. But because both models rely on incorrect beliefs (either over the decision maker’s own preferences or the shocks that they will face) to generate patterns seen in the data, one way to differentiate between the two models is to observe the effects of provision of information about the distribution of costs on subsequent allocative and commitment behavior.<sup>14</sup>

To illustrate the effect of information within the context of the two models, we consider a decision maker who consistently exhibits positive task reallocations and analyze the effect of treating this subject with information (such as the information on past dynamically inconsistent behavior that we treated some subjects with in our experiment) that affects her beliefs.

---

<sup>12</sup>We do not model the source of incorrect beliefs, instead taking them as given and studying their implications. However, a number of existing models could lead to these optimistic beliefs. Kahneman and Tversky (1982) coined the term “planning fallacy” and provided an intuitive model in which decision makers neglect distributional information, leading to optimistic beliefs about outcomes like task duration. Beliefs and updating rules have also been modeled as a choice variable from the point of view of the decision maker (Bénabou and Tirole, 2002, Brunnermeier and Parker, 2005, Brunnermeier et al., 2016). Agents in these models trade off between the distortions caused by incorrect beliefs and their benefits, such as improved self-esteem or higher motivation.

<sup>13</sup>Despite previous work using task reallocations as evidence of bias, the expected value of reallocations can be nonzero even for a decision maker with standard preferences (Strack and Taubinsky, 2021). In general, the expected value depends on the decision maker’s cost function.

<sup>14</sup>Readers interested in one potential formal model consistent with this framework are directed to Breig et al. (2020).



For a naive non-stationary decision maker, we assume that this information causes the decision maker with positive task reallocations to believe they are more present biased. In the notation of O’Donoghue and Rabin (2001), this would imply that  $\hat{\beta}$  falls. Within the context of the two-period decision problem described above, this change in beliefs only affects the decision maker’s willingness to commit—not the allocative choices that she makes when either committed or uncommitted. The effect on commitment demand is straightforward: the decision maker expects that (from the perspective of their current self) her future self will make worse decisions, thus increasing the value of the present self’s choices being implemented. The allocations chosen by the decision maker are not expected to change because the treatment only affects the decision maker’s beliefs, not her preferences: the present self prefers a particular allocation, and changing beliefs about what the future self will do does not change that preference.

**Hypothesis NNS.** For naive non-stationary agents, information provision that increases the perceived level of present bias will increase commitment demand but will have no effect on work allocations.

For the decision maker with biased beliefs about shocks, we assume that information provision makes them more pessimistic about their cost shocks.<sup>15</sup> Such a decision maker modifies her choices so that her earlier decisions are more consistent with the decisions she makes later. Updating beliefs in this way has an ambiguous effect on commitment demand.

**Hypothesis BBS.** For agents with biased beliefs about shocks, information provision that makes beliefs less optimistic will decrease procrastination.

We emphasize that while much of the above discussion is framed in terms of treating a decision maker with systematically positive reallocations (i.e. one who procrastinates), our empirical identification and tests do not rely on the sign of the subject’s reallocation. Subjects who reallocate positively or negatively in this framework are dynamically inconsistent. Informing naïvely nonstationary subjects of either positive or negative reallocation plausibly makes them more likely to believe their preferences are non-stationary. Similarly, for agents with biased beliefs about shocks, information

---

<sup>15</sup>More formally, we assume the updated belief distribution about cost shocks first-order stochastically dominates the prior.

that makes beliefs more optimistic will increase procrastination, while information that makes beliefs more pessimistic will decrease it. Changes in both directions are useful for identifying the source of behavior, and the empirical approach described in Section 3 reflects this.

In what follows, we describe two experiments designed to test these hypotheses. Information in these experiments takes one of two forms. In both experiments, some subjects are provided with reminders about their own past allocative choices. We expect that for naive non-stationary decision makers, reminders about positive (negative) reallocations cause the decision maker to believe that they are more (less) present-biased, increasing (decreasing) commitment demand but leaving allocations unchanged. For decision makers with biased beliefs about shocks, we expect that reminders about positive (negative) reallocations cause the decision maker to be less (more) optimistic about time shocks, decreasing (increasing) reallocations.

In the second experiment, we introduce an experimentally administered cost shock, about which we expect subjects to be optimistic. Subjects compete in a contest, and winning reduces the difficulty of the tasks that they have to complete. Overconfidence in the contest acts similarly to optimism about cost shocks. We provide some subjects with experimentally administered signals relevant to the likelihood that they will win the contest. We hypothesize that negative (positive) signals make decision makers less (more) confident about their likelihood of winning the contest, decreasing (increasing) reallocations.

Because of differences in design between our two experiments, we defer precise discussion of our identification arguments to Sections 3.1.1 and 3.2.1.

## 2.2 Lab-Based Experiment

The hypotheses laid out above were tested using two experiments. The first, lab-based experiment is described here. The experimental instructions and surveys for the lab experiment appear in Appendix I.

To begin, subjects completed an introductory session in the lab.<sup>16</sup> The remainder of the experiment took place during twelve sessions split over two weeks. Table 1 shows a timeline of these sessions. The two weeks of the experiment were identical

---

<sup>16</sup>Subjects were given an overview of the timeline and requirements of the study, completed a survey of basic demographic information as well as a present bias elicitation, did five sample tasks, learned about how the allocation and commitment decisions would be made, and were required to complete a comprehension quiz before advancing.

except for a randomly assigned treatment given during week 2.

The core of the experiment involved subjects making plans for work (real effort tasks) to be carried out Monday and Wednesday of each week, choosing whether to commit to those plans (at a cost in terms of extra tasks), then choosing whether to alter their plans when it came time to actually begin the work on Monday evening. Procrastination was measured by the amount of work that subjects reallocated from Monday to Wednesday relative to their plans.<sup>17</sup> The first week of the experiment was used to gather information on baseline dynamic inconsistency for all subjects. During week 2, a randomly chosen half of the subjects were treated with information on their behavior during week 1. The treatment occurred prior to the subjects making their week 2 work plans. This allows us to study the effect of such information on choices made by the treatment group, testing the hypotheses laid out in Section 2.1, while using the control group to account for anything else that might have changed between weeks 1 and 2.

Table 1: Timeline for Lab Experiment

Monday	Tuesday	Wednesday	Thursday
Morning (6 a.m. – 2 p.m.)			
<ul style="list-style-type: none"> <li>• Bedtime elicitation</li> <li>• <b>(Week 2 only)</b> Randomized treatment</li> <li>• Commitment demand: choose mandatory tasks Mon&amp;Wed</li> <li>• Committed choice: allocate 10 additional tasks across Mon/Wed</li> </ul>	<ul style="list-style-type: none"> <li>• Bedtime elicitation</li> </ul>	<ul style="list-style-type: none"> <li>• Bedtime elicitation</li> </ul>	<ul style="list-style-type: none"> <li>• Bedtime elicitation</li> </ul>
Evening (9 p.m. – 4 a.m.)			
<ul style="list-style-type: none"> <li>• Bedtime plan</li> <li>• Randomized commitment probability &amp; mandatory tasks revealed</li> <li>• Uncommitted choice: allocate 10 additional tasks across Mon/Wed</li> <li>• Complete Mon tasks</li> </ul>		<ul style="list-style-type: none"> <li>• Bedtime plan</li> <li>• Complete Wed tasks</li> </ul>	

Given that the experiment involved dynamic choices, subjects were required to complete surveys and tasks at particular times. A link to each morning survey was sent out at 6 a.m., and subjects were instructed to complete the survey before noon

<sup>17</sup>We also measured dynamic inconsistency around choices of when to go to bed (see Section 2.2.2) in order to test effects on a consequential real-world behavior (Gibson and Shrader, 2018).

that day. At noon, subjects who had not completed the task were sent a reminder and had two hours to complete the survey. If they did not complete the survey by 2 p.m., they were dropped from the study.<sup>18</sup> A link to the evening surveys was sent out at 9 p.m. and the tasks that were part of those surveys had to be completed before 4 a.m. the next morning.

Subjects received \$40 total for completing the full study. An initial payment of \$10 was made to all subjects on Thursday or Friday of the first week. The second payment of \$30 was made to the subjects on Thursday or Friday of the second week, conditional on all portions of the experiment being completed on time.

### 2.2.1 Allocations, Tasks, and Commitment

Subjects made two allocation decisions each week. Each allocation decision consisted of dividing 10 tasks between Monday and Wednesday evenings. On each of these evenings, subjects had to complete the tasks allocated to that evening in addition to a number of mandatory tasks, which are described below. The first allocation was made when completing a survey on Monday morning, imposing at least a seven-hour delay between when the allocative decision was made and when the tasks were actually carried out. The second allocation was made immediately before completing the tasks on Monday evening.

In addition to allocating tasks across evenings, subjects were also offered the chance to commit to their Monday morning choice, increasing the probability that the morning allocation would be the one implemented. If the subjects did not commit they had a one-in-five chance of the morning allocation being implemented. If the subjects did commit this probability rose to four out of five. The commitment was probabilistic rather than deterministic to preserve the incentive compatibility of the evening choices.

To elicit subjects' demand for commitment, they were given the choice of whether or not to commit at a variety of prices, both positive and negative. Due to previous work, including [Augenblick et al. \(2015\)](#), suggesting that many subjects' money-denominated willingness to pay for commitment is near 0, the prices were denominated in terms of mandatory tasks that would have to be done each night in addition to the tasks that were allocated to that night. Mandatory tasks could potentially vary between 4 and 16, depending on a subject's choices and which choice was im-

---

<sup>18</sup>We analyze attrition in Table A1 and find no evidence of selection on observables.

plemented.

The tasks that subjects were required to complete consisted of moving sliders to match particular, predetermined levels. Slider tasks have proved useful in experimental settings as tasks that require real effort and focus from subjects (Gill and Prowse, 2012).<sup>19</sup> A single task consisted of moving nineteen sliders. Each page included no more than 10 tasks. Because each task would fill a computer screen, subjects needed to scroll downward to complete additional tasks. Subjects were unable to proceed to the next page if the current page was incomplete or if there were any errors. If subjects tried to proceed in these cases, they were informed that the task had a problem but were not told which slider was incorrect. The tasks were designed so that each would take about one minute to complete.

### 2.2.2 Bedtime Plans and Actions

Both planned and actual bedtimes were elicited from subjects. In each morning survey, subjects were asked when they went to sleep the night before. Additionally, in both the morning and evening surveys subjects were asked at what time they expected to go to sleep that night. These predictions were deliberately not incentivized because an incentivized prediction could have functioned as a commitment device.<sup>20</sup> The bedtime information allows for tests of changes in real-world time allocation behavior, reported in Breig et al. (2020).

### 2.2.3 Treatment

Within each study wave, randomization was uniform at the subject level.<sup>21</sup> In the second week of the study, treated subjects were given information about their own past choices. The treatment—a real example of which can be seen in Figure 1—consisted of three main parts. The first described the allocation choices that the individual made

---

<sup>19</sup>The required level of each slider was varied to increase difficulty. Each slider was initialized at the number one but had to be clicked before it became active. To avoid subjects becoming confused by their tasks not being accepted due to an inactive slider, the number one was omitted from the potential target levels.

<sup>20</sup>In addition to these self-reported measures, in the first week subjects wore Fitbit wristbands to independently measure their sleep. To obtain the information from the Fitbits, they had to be returned to the experimenter and synced. Due to the time required to sync and recharge each Fitbit, it was infeasible to then immediately return them to the subjects, so the subjects did not have them in week 2 and the experimental treatment focused on survey measures of actual and planned bedtime.

<sup>21</sup>Treatment-control balance is assessed in Table A2. The largest standardized difference between the two groups was that more female subjects were in the treatment group. Results are robust to including or excluding controls for this and other demographic variables.

the week before. Subjects were told whether or not any tasks were reallocated on Monday evening. The second part reported the subject’s average actual and predicted bedtimes and gave the difference between them in minutes. Finally, treated subjects were asked why someone’s choices and predictions might change throughout the day. Subjects were given a blank space in which they had to type something to proceed.

The treatment information was presented neutrally to avoid experimenter demand effects. The message was presented within the survey without an experimenter present, ruling out any physical or vocal suggestions (de Quidt et al., 2019). We provided subjects with information that they could have recorded for themselves had they chosen to do so. Finally, we did not mention commitment.

This information was given to treated subjects (and only treated subjects) on Monday morning of the second week. They were shown the information after they reported their bedtime for the previous night and made a prediction for Monday night but before they made the commitment and allocation decisions.

Figure 1: Treatment

### Choosing the Implemented Allocation

Last week, on Monday morning you said you'd do 15 tasks on Monday evening and 7 tasks on Wednesday. When you were asked in the evening, you decided to do 16 on Monday, and 6 on Wednesday. **Thus, you moved 1 task from Wednesday to Monday.**

Also, on average you predicted that your bedtime would be 12:30 AM, and your actual average bedtime was 1:42 AM, **so you missed your predicted bedtime by about 72 minutes.**

Why might someone's choices and predictions change throughout the day?

There may be unforeseen things that pop up throughout the day that keep them busier than they thought or they miscalculate how long something will take

*Notes:* An example of an actual message that one of the treated subjects received at the beginning of week 2 of the lab experiment, along with the response they entered. The information was provided to subjects just before they made commitment and allocation decisions. The text given in the box is an example of a response that a subject gave to the open-ended question about why someone’s choices and predictions might change. The box was empty when subjects were presented with the message.

## 2.2.4 Sample and Summary Statistics

Undergraduate subjects were recruited to four different sessions of the lab experiment across the second semester of the 2016-2017 academic year. A total of 274 subjects completed the introductory session. Twenty-six of these subjects did not complete some surveys and left the experiment having received only the initial payment of \$10.

The vast majority of those who dropped out of the experiment did so in the first week of their participation. Another 39 subjects missed the completion deadlines for at least one survey, though they eventually did answer all surveys. These subjects are excluded from the primary sample, leaving a final baseline sample of 209 subjects in the lab experiment. Table A1 shows that observable baseline characteristics do not predict attrition. Summary statistics for the final estimation sample are shown in Table A2, and distributions of week 1 and week 2 task reallocation and commitment demand are shown in Figures A1 and A2. Of particular relevance for the estimation results below, the summary statistics show that 35 out of 100 control group subjects and 28 out of 109 treatment group subjects were dynamically inconsistent in week 1 (equally split between those who reallocated positively and negatively).

### 2.3 Online Experiment

The online experiment was pre-registered (Breig et al., 2023) and tested hypotheses similar to those in the lab experiment. It contained one treatment arm that used over-confidence to study biased beliefs about shocks and another that replicated the lab experiment. Full surveys appear in Online Appendix II.

Subjects had to complete five surveys in total. The first survey introduced the experiment.<sup>22</sup> Table 2 provides the timeline for the other surveys, which occurred over two weeks. Like the lab experiment, the online experiment first gathers information on subjects' behavior in week 1 then uses randomized treatments with information in week 2 to study effects on reallocation and commitment demand.

Table 2: Timeline for Online Experiment

Monday/Tuesday	Thursday/Friday
<ul style="list-style-type: none"> <li>•Contest</li> <li>•<b>(Week 2 only)</b> Randomized contest information treatment</li> <li>•Contest belief elicitation</li> <li>•<b>(Week 2 only)</b> Randomized task information treatment</li> <li>•Choose committed allocations</li> <li>•Complete tasks to pay for commitment</li> </ul>	<ul style="list-style-type: none"> <li>•Choose <i>uncommitted</i> allocation <i>unconditional</i> on contest</li> <li>•Learn contest result</li> <li>•Choose <i>uncommitted</i> conditional allocation (easy or hard tasks)</li> <li>•Complete tasks</li> </ul>

<sup>22</sup>In this survey, which always occurred on a Thursday, subjects were informed of the full experimental schedule, completed four example tasks (two easy and two hard, described below), and completed a comprehension quiz about the experiment.



The most important difference between the lab and online experiments is a new treatment based on provision of information about the distribution of costs. In week 1, subjects made allocation decisions (described further below) both over time and over tasks with different levels of difficulty (more or fewer sliders). Subjects were able to do the easy tasks if they won a contest—which consisted of IQ test questions—during the first session of the week. Overconfident subjects would have believed they would be more likely to face the easy tasks, and so would have made more optimistic work plans. During session two of week 1, subjects learned whether they won the contest from the first session. Overconfident subjects would have been more likely than they expected to learn they were facing hard tasks, leading to higher perceived costs and thus reallocation. In week 2, a randomized treatment group was provided with additional information about their contest performance in week 1 prior to making any allocation choices. Specifically, this treated group is told that we matched their performance with two additional randomly drawn contestants, and we tell them the number of times (out of two) that they would have won these matches. We expect this additional information would have led to more accurate beliefs about task difficulty (costs) and less reallocation.<sup>23</sup>

Timing of the surveys was again important to capture dynamic behavior. The second survey was completed on Monday or Tuesday during each study wave. The third survey was completed on Thursday or Friday. Surveys four and five were completed on the following Monday-Tuesday and Thursday-Friday, respectively.<sup>24</sup> Subjects that did not complete a survey by the required time were dropped from the study.<sup>25</sup>

All payments for the online experiment were made the day after the fifth survey of the relevant experiment (on a Saturday). Subjects received a baseline payment of \$9 for completing all five surveys.<sup>26</sup> Given their implemented allocation and piece rate (both described below) they received their payments for completing tasks as a “bonus” on Prolific.

---

<sup>23</sup>Another treatment group received the same type of reallocation behavior information as the treatment group in the lab experiment, allowing for a replication.

<sup>24</sup>The timing restrictions were based on Eastern Standard Time, so subjects on the west coast had to complete surveys 2 and 4 between 9 p.m. on Sunday and 9 p.m. on Tuesday.

<sup>25</sup>We analyze attrition in Table A4.

<sup>26</sup>This amount was calibrated based on pilot data to attain a median hourly wage of \$12/hr for the experiment *excluding* time spent completing the tasks in surveys three and five (which were incentivized through a separate, piece-rate payment).

### 2.3.1 Contests

The contest in the second and fourth surveys was between pairs of participants and involved completing ten questions drawn from the matrix reasoning item bank (MaRs-IB, Chierchia et al. (2019)). Each question contained an incomplete matrix of abstract shapes, with four potential options to complete the matrix. These questions appear in Online Appendix II. Subjects were told that they would complete an “IQ quiz” and that the winner of the contest would receive easier tasks in a future session while the loser would receive more difficult tasks. Subjects were told that in the event of a tie, the winner would be determined randomly.<sup>27</sup> After the contest, we elicited each subject’s beliefs about the likelihood that she had won the contest. These beliefs were elicited without incentivization (Armantier and Treich, 2013).

### 2.3.2 Allocations, Tasks, and Commitment

The allocative decisions that subjects made differed from those in the lab experiment. Rather than splitting a fixed number of tasks between two dates, in surveys two and four subjects were asked to choose how many tasks they would complete in surveys three and five for various piece rates and task difficulty levels. The number of tasks that participants could choose had to be between 0 and 19.<sup>28</sup> The piece rates that subjects faced were \$0.06, \$0.12, and \$0.18 per task.<sup>29</sup> In surveys two and four, subjects made choices conditional on the tasks being easy (winning the contest), being hard (losing the contest), and without knowing the difficulty level (unconditional on contest outcome). In surveys three and five, subjects made choices without knowing the difficulty level (unconditional on contest outcome) and conditional on the *realized* contest outcome (either easy or hard).<sup>30</sup> Thus, subjects made nine allocative choices in surveys two and four and six allocative choices in surveys three and five.

Subjects also made commitment choices in surveys two and four. They chose from a price list that offered an 80% chance of the survey two and four choices being implemented (at the cost of doing one to five easy slider tasks) or an 80% chance

---

<sup>27</sup>In practice, to determine whether a subject won, each subject’s score was recorded and compared against the full distribution of scores of those who completed the quiz in that subject’s wave.

<sup>28</sup>We set the maximum number of tasks to 19 with the hope of avoiding subjects choosing focal round numbers for their allocation.

<sup>29</sup>These piece rates were calibrated based on pilot data to increase the proportion of allocations that were interior.

<sup>30</sup>We expected that making the choice after the contest outcome had been realized would make the realized difficulty level more salient.

of the survey three and five choices being implemented (at the cost of doing three easy slider tasks). Unlike in the lab experiment, the sliders associated with the commitment choice had to be completed at the time commitment was chosen rather than at the same time as those paid by piece rate. This reduced the chances that the randomly drawn commitment price could directly affect the subject’s allocation.

The tasks subjects completed in the online experiment were similar to those from the lab experiment except for the following differences. First, there were two types of tasks: *easy* and *hard*. Each easy task consisted of 20 sliders that needed to be matched to a number between 1 and 20, while each hard task consisted of 30 sliders that needed to be matched to a number between 1 and 30. Second, sliders were initialized at the number zero rather than one. Third, each task needed to be completed on its own page (rather than being grouped into a maximum of 10 tasks per page). Thus, for each task, subjects needed to match all sliders to the relevant number before being able to move on to the next task. Easy tasks were designed so that each would take about one minute to complete, while hard tasks were designed to take between 90 and 120 seconds.

We informed subjects which piece rate they were randomized into immediately after making the allocative choices in surveys three and five. Subjects were then informed whether their committed or uncommitted allocations would be implemented as well as whether their choices that were conditional on contest outcome or unconditional on contest outcome would be implemented. The probability that committed decisions were implemented was either 20% or 80%, and was based on the choice made in the randomly selected row in the commitment price list. The probability of implementing an allocation unconditional on the contest outcome was 50%.

### 2.3.3 Treatments

In the online experiment, treatment occurred during the fourth survey. Subjects were randomized into one of three equal-size treatment groups after entering the survey. The information treatments only modified Survey 4. In the *no information* (control) condition, subjects completed the contest and then proceeded to the belief elicitation and the commitment decision.

In the *contest information* treatment (Figure 2), subjects completed the contest, then received information about their performance in the previous contest. We first reminded the subject of whether they won or lost the first contest and required them

to confirm that they understood this by selecting the appropriate option from “I won the contest in the previous study” and “I did not win the contest in the previous study.” The subject then received information that took the form “We also matched you with two other randomly drawn participants from the previous study, and you (lost against both/won against one/won against both) of them.”<sup>31</sup> They confirmed that they understood the message and then completed the belief elicitation and the commitment decision.

In the *task information* treatment (Figure 3), subjects completed the contest and the belief elicitation, then received information of the form “In Session 2, for a payment rate of (\$0.06/\$0.12/\$0.18) per set and not knowing whether the sets would be easy or hard, you agreed to complete (number chosen) sets. In Session 3, in the same setting, you agreed to complete (number chosen) sets.” The payment rate that was used for each subject was chosen randomly from the three options with equal probability.<sup>32</sup> Subjects in the task information treatment were then asked, “Why might someone’s choices change over time?” They had to type something into a text box as a response to this question, but there were no requirements about what to type other than the box not being empty.

### 2.3.4 Reasoning behind changes between experiments

In this section, we discuss the changes made from the lab experiment to the subsequent online experiment.

First, as discussed in Section 2.2.4, the level of procrastination observed in our student sample was lower than expected: 63 of 209 subjects changed their allocation in the first week, 28 in the treatment group and 35 in the control. Our main object of interest is the *heterogeneous* effect of treatment on subjects based on their week-1 task reallocations. So while the lab experiment had more than 100 subjects in both the treatment and control groups, our main effect is identified by subsamples of sizes 28 and 35. The results that we discuss in Section 3.2.2 show statistical significance using standard inferential procedures, but we conducted a follow-up experiment with a larger sample and altered design to eliminate sample-size concerns.

---

<sup>31</sup>These messages were generated randomly conditional on the subject’s score and the full distribution of scores in the same way as the contest outcome.

<sup>32</sup>We chose to report the *unconditional* allocations because the conditional allocations that subjects made in survey 3 were only for the *realized* contest outcome. Reporting only the unconditional allocations allowed us to randomize which pair of allocations was reported in a way that did not depend on the subject’s performance in the contest.

Figure 2: Contest Information Treatment from the Online Experiment

We already told you that you won the contest from Part 2 of the study. Please confirm that you understand this information

- ☒ I won the contest from the previous study.  
☐ I did not win the contest from the previous study.

We also matched you with two other randomly drawn participants from the previous study, and **you won against one of them**. Please confirm that you understand this information.

- ☐ I won against 0 out of 2 other randomly drawn participants.  
☒ I won against 1 out of 2 other randomly drawn participants.  
☐ I won against 2 out of 2 other randomly drawn participants.

As a reminder, you will win the contest if you have a higher score than the person you are matched with. If you have the same score as the person you are matched with, the winner will be chosen randomly.

**What do you think are the chances, out of 100, that you will win the contest?** You can write down any number from 0 to 100 out of 100.

95

*Notes:* An example of an actual message that one of the subjects in the contest information treatment received in Survey 4 of the online experiment, along with the response they entered. The information was provided to subjects just before they made commitment and allocation decisions. The multiple-choice questions were unselected and the beliefs elicitation was empty when subjects were presented with the message.

A second worry is the causal interpretation of the estimated treatment effect in the lab experiment. While the treatment itself is randomized, our main focus is the *interaction* of the treatment with the level of task reallocation in week 1. Based on our conceptual framework, we expect that this effect arises from the treatment altering the subjects' beliefs about either their naïve non-stationarity or the distribu-

Figure 3: Task Information Treatment from the Online Experiment

In Session 2, for a payment rate of \$0.12 per set and not knowing whether the sets would be easy or hard, you agreed to complete 7 sets. In Session 3, in the same setting, you agreed to complete 3 sets.

So, **the amount of sets you chose in Session 2 is higher than the amount you chose in Session 3.** Please confirm you understand this information.

- ☒ The amount of sets I chose in Session 2 is higher than the amount I chose in Session 3.
- ☐ The amount of sets I chose in Session 2 is the same as the amount I chose in Session 3.
- ☐ The amount of sets I chose in Session 2 is lower than the amount I chose in Session 3.

Why might someone's choices change over time?

I might have felt lazier during Session 3, and not felt the desire to do too many sets for no

*Notes:* An example of an actual message that one of the subjects in the task information treatment received in Survey 4 of the online experiment, along with the response they entered. The information was provided to subjects just before they made commitment and allocation decisions. The multiple-choice question was unselected and the response box was empty when subjects were presented with the message.

tion of time shocks that they face. However, we cannot rule other treatment effects beyond the scope of our model. This concern is addressed with the addition of the contest-information treatment. Because the number of wins of two that is reported to the subject is random *conditional on their score*, the coefficient cannot reflect heterogeneous treatment effects of receiving a message but can only capture the effect of the content of the message. Put differently, in the lab experiment we are interested in understanding the differential effect of receiving the messages “you did more tasks than you originally agreed to” vs. “you did fewer tasks than you originally agreed to.” But in the design of the lab experiment, the populations of subjects that we can send these two messages to are different. In the online experiment, we send subjects with the same contest score (and any other latent characteristics) messages that make

them more or less optimistic.

The information treatments in the online experiment did not involve any information about the subject’s bedtime or bedtime predictions. This allowed for the effects of the “task information” treatment to be measured separately. Because bedtime would no longer be used, the surveys did not include questions about time use or bedtimes, and subjects did not complete more than one survey in a day.<sup>33</sup> The reduction in the number of surveys also allowed for a longer lag between committed allocation decisions and task completion, and the tasks required as payment for the commitment demand decisions could be completed without delay.

The form of the task allocation changed between the lab and online experiments. In the lab experiment, an allocation involved splitting 10 tasks between two days, similar to the design used by [Augenblick et al. \(2015\)](#). In the online experiment, an allocation was a number of tasks that could be completed at various piece rates, with payment occurring at a fixed date after the experiment concluded, similar to the design used in [Augenblick and Rabin \(2018\)](#). We suspect that splitting tasks equally between the two dates was a focal option for subjects in the lab experiment. However, this equal split would also be chosen by time-consistent subjects that did not face time shocks. We believed that paying a piece rate would reduce the chances that subjects appeared time consistent simply because they were choosing focal options.

Finally, the task information treatment in the online experiment should only affect subjects’ beliefs about the costs of completing the tasks rather than the potential benefits in the future. In the lab experiment, one could argue that if the treatment affected subjects’ beliefs about the likelihood of costly time shocks on a Monday night, that it should also affect their beliefs about time shocks on Wednesday night. In the online experiment, the benefit of agreeing to do more tasks is not related to expected business at any future date, so the treatment should only affect beliefs about the distribution of likely costs.

### **2.3.5 Sample and Summary Statistics**

The experiment was completed fully online in two waves during April and May of 2023. Subjects were recruited through Prolific, and all subjects were required to be nationals of the United States. A total of 1,479 subjects completed the introductory session. 1,178 subjects completed all sessions. Following our preanalysis plan, obser-

---

<sup>33</sup>Furthermore, subjects did not receive or use a sleep tracker in the first week.



variations for piece rates under which the subject’s corresponding week-one choice was censored (0 easy tasks or 19 hard tasks) were excluded from the sample to increase statistical power, leaving a final estimation sample of 888 subjects (283 control, 306 treated with task information, 299 treated with contest information).<sup>34</sup> Attrition from the start of the study and after treatment is assessed in Table A4. Observable baseline characteristics do not predict post-treatment attrition. Working as a manager is the variable most strongly associated with overall attrition. Summary statistics for the final estimation sample are shown in Tables A5 and A6, and distributions of week 1 and week 2 task reallocation and commitment demand are shown in Figures A3 and A4. The summary statistics show substantial dynamic inconsistency, with approximately 90% of subjects exhibiting non-zero reallocation for at least one week-1 task choice (committed versus uncommitted).

### 3 Results

#### 3.1 Contest Information Effect on Cost-Based Reallocations

##### 3.1.1 Estimating Equation and Identification

The estimating equation for task reallocation in response to contest information is

$$R_{CD,ir} = \alpha_0 N_{0,i} + \alpha_1 N_{1,i} + \alpha_2 N_{2,i} + \mathbf{Z}_{ir}' \boldsymbol{\theta}_1 + \mathbf{X}_{ir}' \boldsymbol{\theta}_2 + \varepsilon_{ir} \quad (1)$$

where the outcome variable ( $R_{CD,ir}$ ) is the week-2 difference between committed unconditional tasks (difficulty unknown) and uncommitted easy/hard tasks by subject  $i$  for piece rate  $r$ .<sup>35</sup> Whether the uncommitted tasks are easy or hard depends on the outcome of the contest, so the contest provides an observable cost shock. Reallocation  $R_{CD}$  depends in part on this shock, so we call it contest-dependent (CD) reallocation. The primary right-hand-side variables are indicators for whether subjects in the contest information treatment group were told they won zero, one, or two out of two contests against randomly selected competitors ( $N_{0,i}$ ,  $N_{1,i}$ , and  $N_{2,i}$  respectively).

The other variables in the equation are controls and the remaining stochastic error term,  $\varepsilon_{ir}$ .<sup>36</sup> The controls fall into two sets. First are the experimental design controls

---

<sup>34</sup>Empirical results without this sample restriction appear in Tables A7 and A10.

<sup>35</sup>Intuitively, this dependent variable is meant to capture the type of inconsistency observed in real life. Early (committed) choices are made without knowledge of what the final state will be (so they are *unconditional*). Later (uncommitted) choices are made *conditional* on the state, which is observed immediately before taking action.

<sup>36</sup>In all regressions we employ standard errors clustered at the subject level.

$\mathbf{Z}_{ir}$ , which include variables required for conditional exogeneity of contest information and the other, orthogonal treatment that was part of the online experiment (see Section 2.3). The design controls are piece rate fixed effects, indicators for week-one and week-two contest score, indicators for whether the subject won the week-one and week-two contests against her randomly drawn opponent, reported reallocation of tasks in week 1, an indicator for being in the task information treatment group, and the interaction of reported reallocation and the task information treatment group indicator. The second set of controls  $\mathbf{X}_{ir}$  are baseline (week 1) subject characteristics, selected to maximize precision using double machine learning with LASSO (Belloni et al., 2013, Chernozhukov et al., 2018).<sup>37</sup> Equation (1) differs from the pre-analysis plan in minor respects. For this and all other estimating equations, such differences are discussed in Appendix B and evaluated empirically in several appendix tables.

We use Equation (1) to test Hypothesis BBS. As pre-specified, we do so by testing the difference between the estimates  $\hat{\alpha}_2$  and  $\hat{\alpha}_0$ , which measures the difference in week 2 reallocation behavior between a treated subject who was given a signal that she was a relatively strong competitor (two out of two wins) versus a signal that she was a relatively weak competitor (zero out of two wins).<sup>38</sup> A negative signal (zero wins out of two) should make a subject less confident about the likelihood of winning the contest in week 2, leading the subject to believe that she has a higher likelihood of performing hard tasks. The reverse is true for a positive signal (two wins out of two). A positive difference in coefficients provides evidence consistent with BBS, because it implies that subjects who become less optimistic (overconfident) decrease their level of procrastination.

The effects of contest information from Equation (1) are not heterogeneous treatment effects.<sup>39</sup> Because indicators for contest score are included in  $\mathbf{Z}_{ir}$ , the estimated treatment coefficients capture the effect of receiving either positive or negative information, conditional on the underlying probability of winning the contest. Because the information was randomized within groups with the same score, these groups were balanced in expectation on both unobservable and observable characteristics.

---

<sup>37</sup>Results without these controls are presented in Table A8 and show similar point estimates. The set of possible variables for the LASSO procedure are the difference between committed unconditional tasks and uncommitted easy/hard tasks in week 1 as well as indicators for experiment wave, education level, employment status, elicited risk tolerance, elicited patience, age, gender, ethnicity, country of birth, languages spoken, student status, and working as a manager.

<sup>38</sup>The effect of the one-win message ( $\alpha_1$ ) is ex ante ambiguous and we do not interpret it.

<sup>39</sup>That is, randomized treatment variables do not interact with endogenous subject characteristics.

### 3.1.2 Overconfidence Online

Before examining contest-dependent task reallocation, we first present evidence on the mechanisms discussed in Section 2.1. Hypothesis BBS asserts that contest information affects beliefs, making earlier decisions more consistent with later decisions. That is, beliefs change, so plans change, and the end result is a change in reallocation. To assess empirically the first two links in this hypothesized causal chain, we estimate contest-information treatment effects on beliefs about winning the contest and earlier decisions (committed allocations).

Table 3: Effects of contest information on beliefs, weight on easy allocation

	(1) Win belief	(2) $w_e$
0-win message ( $\hat{\alpha}_0$ )	-11.3 (2.93)	-0.19 (0.090)
1-win message ( $\hat{\alpha}_1$ )	-1.86 (2.58)	-0.032 (0.072)
2-win message ( $\hat{\alpha}_2$ )	4.80 (2.34)	0.052 (0.068)
$\hat{\alpha}_2 - \hat{\alpha}_0$	16.1	0.24
Right-tailed $p$ value	8.0e-07	.012
Subjects	888	565
Observations	888	1301

*Notes:* Both columns shows results from estimating versions of equation (1) on the online experiment sample. The dependent variable in Column (1) is the elicited probability that a subject would win the contest in week 2. In Column (2) it is the weight  $w_e$  on the committed easy allocation. The sample size is reduced, as  $w_e$  is undefined for subjects whose committed easy and committed hard allocations are identical. In parentheses are standard errors clustered at the subject level.

Effects on beliefs about winning the contest appear in Table 3 Column 1. The results are based on estimating a subject-level version of equation (1) where the dependent variable is the probability that the subjects reported for whether they would win the contest in week 2.<sup>40</sup> Because priors are necessarily on the  $[0, 100]$

<sup>40</sup>The regression omits controls for week 2 contest score or the week 2 win indicator because those were not known by the subjects at the time they reported beliefs.

interval, a zero-win message should weakly decrease win belief and a two-win message should weakly increase it.<sup>41</sup> Estimates are consistent with these predictions. On average, subjects who received a zero-win message decreased their win beliefs by approximately 11 percentage points, while subjects who received a two-win message increased win beliefs by 4.8 percentage points. In relative terms, the two-win message led to a 16 percentage point increase in reported probability of winning compared to the zero-win message (proportionally, a 30% increase on over average week 1 beliefs).

Table 3 Column 2 evaluates the second link in the chain: whether the contest information that changed beliefs also changed plans. We refer to the dependent variable of this regression as the *weight on the committed easy allocation*, denoted  $w_e$ . It is defined as the difference between the committed unconditional and committed hard allocations divided by the difference between the committed easy and committed hard allocations.<sup>42</sup> Intuitively,  $w_e$  is the distance between the unconditional and hard allocations, expressed as a proportion of the distance between the easy and hard allocations. For example, a high  $w_e$  means the unconditional allocation is far from the hard allocation, and therefore close to the easy allocation. The results in Table 3 Column 2 show that the 2-win message brings the unconditional allocation closer to the easy allocation, while the 0-win message brings the unconditional allocation closer to the hard allocation. This is consistent with the logic of Section 2.1: when subjects become less optimistic, their unconditional choices move closer to what they choose when they *know* the bad outcome will occur.

Having evaluated mechanisms, we proceed to the outcome of primary interest: reallocation behavior. Table 4 shows the results of estimating equation (1) where the dependent variable is contest-dependent reallocation ( $R_{CD}$ ). Treated individuals who received a zero-win message decreased their contest-dependent task reallocations relative to the control group, while two-win message recipients did the opposite. Comparing the effects on these two groups (taking the difference between  $\hat{\alpha}_2$  and  $\hat{\alpha}_0$ ), the gap between planned and actual work was one task greater for subjects who received the two-win message relative to the those who received the zero-win message

---

<sup>41</sup>The effect of a one-win message is theoretically ambiguous and our pre-specified test does not involve this coefficient.

<sup>42</sup>In other words,  $w_e$  is found by solving  $u = w_e e - (1 - w_e)h$  where  $u$  is the committed unconditional allocation,  $e$  is the committed easy allocation, and  $h$  is the committed hard allocation. For subjects who set the easy allocation equal to the hard allocation at a given piece rate, this ratio is undefined, and such observations are excluded from the regression. Table A13 reports similar effects on related dependent variables,  $\frac{w_e}{1 - w_e}$  and  $u - h$  (the numerator of  $w_e$ , which is defined for all subjects).

Table 4: Effect of contest information on contest-dependent task reallocation

	$R_{CD}$ reallocation
0-win message ( $\hat{\alpha}_0$ )	-0.20 (0.30)
1-win message ( $\hat{\alpha}_1$ )	0.43 (0.28)
2-win message ( $\hat{\alpha}_2$ )	0.88 (0.32)
$\hat{\alpha}_2 - \hat{\alpha}_0$	1.08
Right-tailed $p$ value	.0036
Subjects	888
Observations	2322

*Notes:* Results are from estimating equation (1) using the online experiment sample. The dependent variable is contest-dependent reallocation (committed unconditional minus uncommitted easy/hard). In parentheses are standard errors clustered at the subject level.

(right-tailed  $p$  value of 0.0036).<sup>43</sup>

This result is clearly consistent with Hypothesis BBS. Information that makes subjects less optimistic (0-win signals) decreases reallocations, while information that makes them more optimistic (2-win signals) increases reallocations. Because this information has an effect on allocations, the result is not consistent with Hypothesis NNS, showing that reallocations cannot be caused solely by naïve nonstationarity.

### 3.2 Procrastination Reminder Effect on Time-Based Reallocation and Commitment Demand

#### 3.2.1 Estimating Equation and Identification

In this section, we examine the effects of information about past procrastination on both task reallocation and commitment demand, in both the lab and online experiments. The estimating equations for all results are versions of the following:

$$y_i = \gamma_1 W_i + \gamma_2 R_{R,i1} + \gamma_3 W_i R_{R,i1} + \mathbf{Z}_i' \boldsymbol{\theta}_3 + \mathbf{X}_i' \boldsymbol{\theta}_4 + \nu_i. \quad (2)$$

<sup>43</sup>The test statistic  $\hat{\alpha}_2 - \hat{\alpha}_0$  and the one-tailed alternative hypothesis were pre-specified.

The dependent variable, denoted generically by  $y_i$ , depends on the hypothesis being studied. To study commitment demand in both the online and lab-based experiments, the outcome variable is the change in commitment demand for subject  $i$  between week 1 and week 2, denoted  $C_i$ . To study task reallocation, the outcome variable is the number of tasks reallocated in week 2. For the lab experiment, this is the number of tasks that the subject committed to completing minus the uncommitted allocation that the subject completed. For the online experiment, this is  $R_{CI,ir}$ , the difference between committed and uncommitted allocations within piece rate and information condition (easy/hard/unconditional). This reallocation is contest-independent (CI), in contrast to the contest-dependent reallocation ( $R_{CD,ir}$ ) analyzed in Section 3.1.

The primary right-hand-side variables are an indicator for being in the reallocation reminder treatment group ( $W_i$ ), task reallocation in week 1 (which we denote  $R_{R,i1}$  because it is the reallocation behavior that was *reported* to treated subjects), and the interaction of these two variables.<sup>44</sup> The vector  $\mathbf{Z}_i$  contains control variables for features of the experimental design and the vector  $\mathbf{X}_i$  contains controls for baseline (week 1) characteristics,<sup>45</sup> while  $\nu_i$  is the stochastic error term.

The main parameter of interest is  $\gamma_3$ , which we use to test Hypotheses BBS and NNS. The hypotheses concern the effect on subsequent choices of changes in beliefs, which we assume are affected by information on past choices. For naive non-stationary subjects, we assume that reminders about past choices affect beliefs about the subject’s own preferences. For optimistic subjects, we assume that reminders about past choices affect beliefs about future shocks. As we control for week-1 reallocation,  $\gamma_3$  captures whether information-treated subjects reallocate differently in week 2 than their week-1 behavior would predict.

One important additional assumption must be invoked for the lab experiment: that for optimistic subjects, the treatment affected the subjects’ beliefs more about night 1 (Monday) than night 2 (Wednesday). This is a natural assumption because the treatment reminded the subjects specifically about choices made on Monday of

---

<sup>44</sup>In the online experiment task reallocation regressions, we randomly selected one of the piece rates and provided information on that to the treated subjects. The outcome variable in that specification is piece-rate specific, and subjects could contribute up to three observations to the regression.

<sup>45</sup>For the lab experiment, design controls are indicators for study wave and receipt of a survey completion reminder in week 1. The set of possible variables for the LASSO procedure are gender and age indicators, GPA and GPA squared, an employment indicator, and week-one time spent on socializing and studying. For the online experiment, controls are identical to those described in Section 3.1.1. Results without LASSO-selected controls appear in Tables A3 and A11.

the previous week.<sup>46</sup> This assumption is not required for the online experiment.<sup>47</sup>

### 3.2.2 Optimism and Present Bias in the Lab

Table 5 shows the results from reporting past reallocation behavior. Column 1 shows the effect on week 2 task reallocation, and the additional effect of being treated with non-zero information is shown by coefficient on the interaction. The estimate of -0.44 shows that for subjects who reallocated work in week 1, the treatment caused a statistically significant and practically substantial reduction in the reallocation of tasks in week 2.<sup>48</sup> Column 2 shows the treatment effect on the change in commitment demand, with the estimated coefficient on the interaction term positive and statistically significant. This estimate is consistent with a present-biased individual updating her belief over her present bias and increasing her commitment demand in response.

Table 5: Lab experiment: Effect of task information on contest-independent task reallocation and commitment demand

	(1) $R_{CI}$ reallocation	(2) $\Delta$ Commitment demand
Task message	0.20 (0.37)	0.040 (0.41)
Reported reallocation	0.31 (0.13)	0.015 (0.10)
Task message $\times$ reported reallocation	-0.44 (0.20)	0.44 (0.18)
Two-tailed $p$ value, interaction	0.027	0.011
Subjects	209	209
Observations	209	209

*Notes:* The table shows results from estimating equation (2) on the lab experiment sample. Column (1) shows the effect on task reallocation of being treated with messages about week 1 task reallocation. Column (2) shows the effect of the same treatment on the change in commitment demand. In parentheses are standard errors clustered at the subject level.

The estimates from Table 5 Column 1 are inconsistent with Hypothesis NNS,

<sup>46</sup>Formally, this will hold if the distribution of cost shocks depends partly on the day of the week.

<sup>47</sup>All tasks in the online experiment were completed at once at a piece rate. Information about time shocks could only affect beliefs about costs, not about the benefits of completing more tasks.

<sup>48</sup>The other coefficients are not of primary interest. The coefficient on reallocation in week 1 shows that a subject's reallocations are correlated across weeks: Subjects who reallocated more tasks in week 1 also did so in week 2.



but are consistent with Hypothesis BBS. Intuitively, our test is whether reminding a subject of a past reallocation causes them to update their beliefs either about their own non-stationarity or about the shocks that they face. The fact that reallocations change in week 2 as a result of these reminders indicates that the information cannot *solely* be affecting beliefs about nonstationarity.

However, the evidence from Table 5 Column 2 shows that the reminders did have an effect on commitment demand, consistent with Hypothesis NNS. Our interpretation is that naive non-stationarity contributes to the reallocations of subjects in the lab experiment.<sup>49</sup>

### 3.2.3 Optimism and Present Bias Online

Table 6 reports results from the online replication of the lab-based experiment. The structure of the table follows that of Table 5, with the first column showing the effect of treatment on contest-independent task reallocation behavior and the second column showing the effect on the change in commitment demand.

Table 6: Online experiment: Effect of task information on contest-independent task reallocation and commitment demand

	(1) $R_{CI}$ reallocation	(2) $\Delta$ Commitment demand
Task message	0.31 (0.14)	0.075 (0.11)
Reported reallocation	0.10 (0.021)	-0.0052 (0.015)
Task message $\times$ reported reallocation	-0.062 (0.037)	0.0072 (0.025)
Left-tailed $p$ value, interaction	0.045	0.39
Subjects	888	888
Observations	4644	888

*Notes:* Results are from estimating equation (2) on the online experiment sample. Column (1) shows the effect on contest-independent task reallocation of being treated with messages about week 1 reported task reallocation. Column (2) shows the effect of the same treatment on the change in commitment demand. In parentheses are standard errors clustered at the subject level.

<sup>49</sup>While estimates could be consistent with updating the beliefs of an agent with biased beliefs about shocks, the theoretical prediction in that context is ambiguous and we surmise that any such effect would be small.

The effect of the interaction of treatment on individuals with a larger reported reallocation is qualitatively similar to the result from the lab experiment. The treatment caused these individuals to reduce their task-independent reallocation in week 2 by 0.06.<sup>50</sup> As in the lab experiment (Table 5 Column 1), these results are inconsistent with Hypothesis NNS but consistent with Hypothesis BBS.

Table 6 Column 2 shows that, in contrast to the lab experiment, treatment did not have a statistically significant effect on the change in commitment demand for individuals who reallocated more in week 1. There are multiple potential explanations. First, sample populations differed. Students might be more present biased or more likely to update about their bias. While there do appear to be differences in the initial distributions of commitment demand between samples, we do not find substantially different average reallocation behavior. Second, there were slight differences between the experiments in the way commitment demand was elicited. The range of possible “prices” (denominated in tasks) was smaller online, and the tasks that the subjects had to do to pay for commitment were completed at the time of commitment rather than alongside the other tasks. Finally, it is possible that the original result was a false positive.

## 4 Conclusion

This paper models agents whose dynamic inconsistency potentially arises from two sources: naive non-stationary preferences and biased beliefs over cost shocks. Agents with optimistic beliefs about shocks will exhibit dynamically-inconsistent choices over effort that are observationally equivalent to those driven by naive non-stationarity. An information intervention that tells agents about their past time inconsistency, however, can distinguish these models: optimistic agents will change effort allocations, but naive agents with non-stationary preferences will not.

We test these predictions experimentally and find that biased beliefs about shocks do matter for time inconsistency. The results help explain puzzlingly low take-up of costly commitment. Perhaps more importantly, they offer an alternative policy prescription to help overcome time-inconsistent behavior—providing information on agents’ own past execution of their plans just prior to a new decision. One avenue for

---

<sup>50</sup>Comparing the results in the lab and online experiments, we find that among untreated subjects, an additional unit of reallocation in Week 1 is associated with an additional Week 2 reallocation of 0.3 tasks in the lab experiment, but only 0.1 tasks in the online experiment. Thus it is not surprising that the magnitude of the treatment effect also falls.

future research is to identify situations in which subjects demand this information, and how it can be structured to reduce time inconsistency with minimal associated welfare losses.

## References

- Acland, D. and M. R. Levy (2015). Naiveté, projection bias, and habit formation in gym attendance. *Management Science* 61(1), 146–160.
- Allcott, H., J. J. Kim, D. Taubinsky, and J. Zinman (2021). Are high-interest loans predatory? theory and evidence from payday lending. pp. 1–58.
- Andreoni, J. and C. Sprenger (2012). Risk Preferences Are Not Time Preferences. *American Economic Review* 102(7), 3357–3376.
- Armantier, O. and N. Treich (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review* 62, 17–40.
- Ashraf, N., D. Karlan, and W. Yin (2006). Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *Quarterly Journal of Economics* 121(2), 635–672.
- Augenblick, N., M. Niederle, and C. Sprenger (2015). Working over time: Dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics* (2001), 1067–1115.
- Augenblick, N. and M. Rabin (2018). An experiment on time preference and misprediction in unpleasant tasks. *Review of Economic Studies*.
- Barro, R. J. (1999). Ramsey meets Laibson in the neoclassical growth model. *Quarterly Journal of Economics* 114(4), 1125–1152.
- Belloni, A., V. Chernozhukov, and C. Hansen (2013). Inference on Treatment Effects after Selection among High-Dimensional Controls. *The Review of Economic Studies* 81(2), 608–650.
- Bénabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics* 117(3), 871–915.
- Börsch-Supan, A. H., T. Bucher-Koenen, M. D. Hurd, and S. Rohwedder (2018). Saving regret.
- Breig, Z., M. Gibson, and J. Shrader (2020). Why do we procrastinate? Present bias and optimism. *IZA Working Paper No. 13060*.
- Breig, Z., M. Gibson, and J. Shrader (2023). Why do people change their work plans? *AEA RCT Registry*.
- Brunnermeier, M. K., F. Papakonstantinou, and J. A. Parker (2016). Optimal time-inconsistent beliefs: Misplanning, procrastination, and commitment. *Management Science* 63(5), 1318–1340.
- Brunnermeier, M. K. and J. A. Parker (2005). Optimal expectations. *American Economic Review* 95(4), 1092–1118.

- Chakraborty, A. (2021). Present bias. *Econometrica* 89(4), 1921–1961.
- Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal* 21(1), C1–C68.
- Chierchia, G., D. Fuhrmann, L. J. Knoll, B. P. Pi-Sunyer, A. L. Sakhardande, and S.-J. Blakemore (2019). The matrix reasoning item bank (mars-ib): novel, open-access abstract reasoning items for adolescents and adults. *Royal Society open science* 6(10), 190232.
- de Quidt, J., L. Vesterlund, and A. J. Wilson (2019). Experimenter demand effects. In *Handbook of Research Methods and Applications in Experimental Economics*. Edward Elgar Publishing.
- DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature* 47(2), 315–372.
- DellaVigna, S. and U. Malmendier (2006). Paying not to go to the gym. *American Economic Review* 96(3), 694–719.
- Eyal, P., R. David, G. Andrew, E. Zak, and D. Ekaterina (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 1–20.
- Gibson, M. and J. Shrader (2018). Time use and labor productivity: The returns to sleep. *Review of Economics and Statistics* 100, 783–798.
- Gill, D. and V. Prowse (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review* 102(1), 469–503.
- Halevy, Y. (2008). Strotz meets Allais: Diminishing impatience and the certainty effect. *American Economic Review* 98(3), 1145–62.
- Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica* 83(1), 335–352.
- Heidhues, P. and B. Köszegi (2010). Exploiting naivete about self-control in the credit market. *American Economic Review* 100(5), 2279–2303.
- Kahneman, D. and A. Tversky (1982). Intuitive prediction: Biases and corrective procedures. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*, pp. 414–421. Cambridge University Press.
- Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society*, 287–309.
- Krugman, P. (2015). Wearables and self-awareness. <https://krugman.blogs.nytimes.com/2015/03/09/wearables-and-self-awareness-personal/>.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112(2), 443–478.
- Laibson, D. (2015). Why Don’t Present-Biased Agents Make Commitments? *American Economic Review* 105(5), 267–272.

- Le Yaouanq, Y. and P. Schwardmann (2022). Learning about one’s self. *Journal of the European Economic Association* 20(5), 1791–1828.
- O’Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review* 89(1), 103–124.
- O’Donoghue, T. and M. Rabin (2001). Choice and procrastination. *The Quarterly Journal of Economics* 116(1), 121–160.
- Roy, M. M., N. J. Christenfeld, and C. R. McKenzie (2005). Underestimating the duration of future events: Memory incorrectly used or memory bias? *Psychological Bulletin* 131(5), 738.
- Schilbach, F. (2019). Alcohol and Self-Control: A Field Experiment in India. *American Economic Review* 109(4), 1290–1322.
- Strack, P. and D. Taubinsky (2021). Dynamic preference “reversals” and time inconsistency.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies* 23(3), 165.
- Toussaert, S. (2018). Eliciting temptation and self-control through menu choices: A lab experiment. *Econometrica* 86(3), 859–889.