Asymmetric Naïveté: Beliefs about Self-Control

Anastassia Fedyk*

August 27, 2018

Abstract

Do individuals anticipate present bias in others? This paper jointly investigates beliefs about one's own and others' present bias. In an online experiment, participants engaged in a real-effort task display little awareness of their own present bias, but anticipate present bias in others. Structurally, I estimate a present bias parameter β of 0.82. Participants perceive others' β to be 0.87, indicating substantial sophistication, contrasted with 1.03 for themselves, indicating full naïveté. A field survey in the classroom further confirms that individuals appear to be aware of present bias in others, but are overoptimistic about themselves.

Keywords: time preference, beliefs, overconfidence

*UC Berkeley, Haas School of Business. Mail: 545 Student Services Building #1900, Berkeley, CA 94720. Email: fedyk@berkeley.edu. I am grateful to Philippe Aghion, Ned Augenblick, Linda Babcock, Max Bazerman, John Beshears, Vitaly Bord, Kirill Borusyak, John Campbell, Katie Coffman, Lucas Coffman, Lauren Cohen, Stefano DellaVigna, Benjamin Enke, Jerry Green, Oliver Hart, James Hodson, Xavier Jaravel, Marc Kaufmann, David Laibson, Andres Liberman, George Loewenstein, Michael Luca, Kristof Madarasz, Chris Malloy, Emily Oster, Matthew Rabin, Alex Rees-Jones, Todd Rogers, Andrei Shleifer, Charles Sprenger, Jeremy Stein, Tomasz Strzalecki, Dmitry Taubinsky, Florian Zimmermann, and seminar participants at Harvard University, the 2017 American Economic Association, the 2017 SITE Workshop on Experimental Economics, and the 2018 SITE Workshop on Psychology and Economics for helpful comments. I am thankful to Tatiana Fedyk for administering the classroom experiment in her class at the University of San Francisco. Funding for this project was generously provided by the Harvard Business School, the Lab for Economic Applications and Policy, the Pershing Square Venture Fund for Research on the Foundations of Human Behavior, and the Russell Sage Foundation. The research in this paper was approved by the Harvard University IRB. All remaining errors are my own.

1 Introduction

While time-inconsistent preferences have been gaining prominence in economics, helping explain a variety of individual behaviors ranging from life-time savings to exit rates from unemployment,[1] empirical work has concentrated on individuals' awareness of their own present bias without considering beliefs regarding others.[2] Yet many situations where time-inconsistent preferences are likely to play a key role (including teams in corporations, households' consumption decisions, and educational settings) involve *interactions* among biased individuals. Households' savings decisions and demand for commitment devices depend on the spouses' expectations regarding each others' future behavior. In the workplace, managers' ability to effectively delegate tasks hinges on their awareness of their subordinates' present bias. And across both educational and corporate environments, efficacy of incentive schemes such as tournaments depends on the individuals' relative perception of themselves compared to others. In these situations, understanding of *others'* present bias plays a key role in determining equilibrium outcomes.

In order to lay the foundation for analyzing interactions among present biased individuals, this paper experimentally investigates individuals' beliefs regarding their own and others' present bias within a single unified framework. To what extent are people aware of the self-control problems of others? Are beliefs regarding others more correct than those regarding self? I document that while individuals are largely unaware of their own tendency to procrastinate, they hold much more sophisticated beliefs about others. This wedge in beliefs is consistent with the notion of bias blind spots documented in the social psychology literature,[3] and suggests that naïveté regarding one's own present bias is a form of overconfidence rather than a lack of awareness of time-inconsistency in general.

I measure beliefs regarding one's own and others' present bias using both laboratory experiment and field survey evidence. First, I construct a large-scale online laboratory experiment to isolate these beliefs. The experiment addresses the issue of incentive compatibility and allows for structural estimation of parameters reflecting beliefs regarding self and others. The results point to a wedge in beliefs: the participants are naïve about their own present bias, but expect present bias in others. Second, as a test of the external validity of the documented wedge in beliefs, I conduct a field survey in an undergraduate accounting class. The classroom experiment confirms that the wedge in beliefs between self and others

[1]See, for example, Laibson (1997), DellaVigna and Paserman (2005), and Laibson, Repetto, and Tobacman (2008).

[2]See, for example, Ariely and Wertenbroch (2002), DellaVigna and Malmendier (2006), Acland and Levy (2015), Augenblick, Niederle, and Sprenger (2015), and Augenblick and Rabin (2018).

[3]See Pronin, Lin, and Ross (2002), Ehrlinger, Gilovich, and Ross (2005), and West, Meserve, and Stanovich (2012).

is operative and substantial in a real-world setting – the classroom.

The online laboratory experiment runs over the course of four weeks, and recruits participants from the Harvard Decision Sciences Lab. The participants engage in a real-effort task that involves identifying characters on a computer screen, and are asked how much work they would like to perform at different wages. Work decisions are elicited for the current date and for future dates, allowing for an estimate of present bias to be computed by comparing decisions about future work to decisions about immediate work. Some of the participants are also asked, on each date, to predict the choices that they will make on future dates about immediate work on those dates. This provides an estimate of the degree to which the participants are aware of the time-inconsistency in their own preferences. Another group of participants is asked to predict the average answers of the other experimental participants – both the others' current preferences for future dates and the choices that others will make when the future dates actually arrive. The predicted differences capture the beliefs that experimental participants hold about others' present bias. To investigate the robustness of the results to asking the two sets of prediction questions (about self and about others) separately versus together, a third group of participants faces both sets of questions.

My experimental design introduces three innovations relative to previous online experiments on present bias using real effort tasks (Augenblick, Niederle, and Sprenger, 2015; Augenblick and Rabin, 2017). First, in order to elicit participants' beliefs regarding others and not just themselves, I take steps to ensure that experimental participants are correctly calibrated regarding the population of "others" for their predictions. I do so through an interactive display of the demographics and self-reported task-enjoyment from the pilot run of the experiment.[4] The interactive display provides summaries by gender, age, race, marital status, educational attainment, and employment. Second, given that the real effort task is performed online, it is important to ensure that "immediate" work decisions are indeed perceived as imminent.[5] On each date, as soon as a work decision is selected to be implemented, the participant must complete the chosen amount of work immediately, with a total of no more than fifteen minutes of break. A prominently displayed timer on the webpage alerts the participant to the countdown. Third, since laboratory subject pools at universities tend to consist of homogeneous populations largely featuring students, I implement a stag-

[4]Importantly, the pilot draws from the same Harvard Decision Sciences Lab participant pool, and the composition of participants does not differ across the pilot and the main experiment.

[5]Several papers address the degree of immediacy necessary for present bias to be operative, especially centered around whether immediate monetary rewards translate into immediate consumption. See, for example, Augenblick, Niederle, and Sprenger (2015) and Balakrishnan, Haushofer, and Jakiela (2017). In online experiments, one concern is whether participants' choices of work on a given date are truly immediate, if, for example, the choice is made in the morning but the participant then leaves the experimental screen inactive until the evening.

gered sessions design to ensure that the participants' answers are not affected by systematic shocks such as school deadlines, university-wide events, or weather disruptions. I run the experiment in five non-overlapping sessions spread across January - August, 2016.

The results of the experiment indicate that participants display significant present bias, are quite naïve about their own present bias, and are more aware of present bias in others. The participants in the online work experiment choose, on average, 3.50 rounds of work fewer when faced with decisions that have immediate consequences than when they make the decisions ahead of time. There is virtually no self-awareness regarding this time-inconsistency in preferences when participants are asked to predict their own future decisions. However, when asked to predict the decisions of others, participants expect others to choose an average of 1.49 rounds fewer when the choice is made for immediate work than when the choice is made for future work. The results are robust to posing the self- and other-prediction questions separately across participants and together to the same participants.

I exploit the rich and controlled setting of my experimental design to structurally estimate the extent of the participants' present bias, naïveté, and beliefs regarding others. I consider a standard β-δ model of quasi-hyperbolic discounting,[6] coupled with a separable utility function consisting of a linear utility in money and a power cost of effort, and allow for a misunderstanding of the present bias parameter β when participants predict their own or others' future decisions. I use the participants' decisions and predictions for different dates at different wages to estimate the model's parameters.

Pooling across all of the participants' work decisions, I document a present bias parameter β of approximately 0.82, which is consistent with prior literature.[7] The participants' self-predictions reveal no awareness of their own present bias: on average, they perceive their present bias parameter to be 1.03. By contrast, predictions regarding other participants indicate strong, albeit incomplete, awareness of others' present bias: participants perceive others' β to be around 0.87, higher than the true value of 0.82, but statistically significantly different from 1. These estimates are robust to excluding participants who do not complete the entire experiment, to allowing for different predicted baseline levels for self versus others, and to estimating the model separately for answers about self and others. The results are also corroborated by estimating the model individually for each participant. The median value of the present bias parameter β across individual participants is 0.92. The median self-prediction is at 1.00, and the median prediction for others is at 0.93.

To illustrate the practical relevance of the experimentally documented wedge in beliefs in

[6] See, for example, Laibson (1997), O'Donoghue and Rabin (1999a), DellaVigna and Paserman (2005), and Heidhues and Kőszegi (2010).

[7] For example, Shui and Ausubel (2005) estimate β around 0.8, Laibson et al. (2008) estimate β at 0.71, and Augenblick and Rabin (2018) estimate β at 0.83.

a field setting, I run a secondary experiment in an undergraduate financial accounting class at the University of San Francisco. The students are assigned an Individual Project, which requires them to choose a publicly traded company and analyze its financial statements by May 2, 2016. The students must choose a company to analyze, download its financial statements ahead of time, and email their selection for instructor approval by April 2, 2016. There is an incentive for completing this part of the assignment earlier: no two students can analyze the same company, and the companies are allocated on a first-come-first-served basis. On the first day of class, January 25, 2016, a survey is administered to the class, asking students to predict when they and / or their average classmates would submit their selection for instructor approval. The voluntary and fully anonymous survey has a randomized structure with three arms: (i) students are asked to predict the average date when they will email their selection to the instructor; or (ii) students are asked to predict the date when their classmates will email their selections to the instructor; or (iii) students are asked to make both predictions.

The results of the classroom experiment confirm the students' naïveté about their own procrastination coupled with more sophisticated beliefs about others. The students predict that they will send their chosen company to the instructor, on average, 22 days before the deadline. By contrast, the students expect that their peers will email the instructor an average of 9 days before the deadline. The actual dates when the students email the instructor occur, on average, 7 days before the deadline, indicating that the predictions for self are optimistic, while the predictions for others are well calibrated (the average prediction for others is not statistically different from the average actual submission date). The difference in predictions for self and others is highly statistically significant and robust to posing the self- and other- predicting questions separately to different students or together to the same students.

The classroom experiment demonstrates the relevance of asymmetric naïveté in one real-world setting, but the wedge in beliefs can also influence equilibrium outcomes across a wide spectrum of competitive, collaborative, and hierarchical environments. Relative performance metrics and tournament incentives are ubiquitous both in the workplace (e.g., bonuses for top performance) and in the classroom (e.g., grading on a curve). Across these situations, an individual's willingness to enter a tournament incentive scheme and her subsequent level of effort depend on her expectations regarding her peers' behavior. Similarly, in collaborative environments such as households or teams of coworkers, willingness to enter into commitment devices such as deadlines or savings contracts depends on each individual's perception of both her own and her partners' present bias. Lastly, in hierarchical settings, a teacher's ability to optimally structure class assignments hinges on his understanding of his students' present bias, while a manager's effectiveness at delegating, structuring tasks, and setting deadlines depends on her beliefs regarding her employees. A more detailed review of the extensive applications of asymmetric naïveté across competitive, collaborative, and hierarchical settings is presented in Section 6.

This paper contributes to the growing experimental literature on time preference and naïveté. Multiple prior studies experimentally assess the extent of individuals' present bias[8] and awareness of their own time-inconsistency.[9] These studies document present bias in the domains of monetary rewards, food choice, and real effort, and find a fair amount of naïveté regarding one's own present bias. The present paper extends this line of work by jointly investigating beliefs about self and beliefs about others, and the extent to which the previously documented naïveté is a systematic underestimation of present bias in general or optimism specifically about one's own self-control. I offer experimental evidence in favor of the latter: individuals are generally aware of present bias in others, and are overoptimistic specifically about themselves.

The results on the wedge in beliefs also lay the foundations for theoretical studies of interactions between biased agents. Naïveté regarding one's own present bias has informed a number of theoretical works, including DellaVigna and Malmendier (2004) and Heidhues and Kőszegi (2010).[10] While beliefs regarding self suffice for these models of single biased agents and rational principals, recent theoretical studies have begun to model interactions between multiple present-biased agents. For example, Fahn and Hakenes (2014) consider fully sophisticated agents who are aware of their own and others' self-control problems. Fedyk (2015) assumes that agents are at least partially naïve about their own present bias, but hold more accurate beliefs regarding others. The investigation of individuals' awareness of others' present bias will serve to ground models of interactions between present-biased individuals with experimentally-tested assumptions.

The remainder of the paper proceeds as follows. Section 2 outlines the design of the online laboratory experiment, while Section 3 presents the reduced-form results. Section 4 presents the structural model and estimates the belief parameters for self and others. Section 5 presents field evidence of the wedge in beliefs regarding own and others' procrastination in the classroom. Section 6 discusses applications of the documented wedge in beliefs. Section 7 concludes.

[8]See Solnick, Kannenberg, Eckerman, and Waller (1980), Read and van Leeuwen (1998), McClure, Laibson, Loewenstein, and Cohen (2004), Andersen, Harrison, Lau, and Rutstrom (2008), Tanaka, Camerer, and Nguyen (2010), Bisin and Hyndman (2014), among others.

[9]See Ariely and Wertenbroch (2002), DellaVigna and Malmendier (2006), Skiba and Tobacman (2009), Acland and Levy (2015), and Augenblick and Rabin (2018).

[10]See also O'Donoghue and Rabin (1999a, 1999b), Gul and Pesendorfer (2001), Gottlieb (2008), and Herweg and Müller (2011), among others.

2 Experimental Design

In this section, I detail the design of the online laboratory experiment used to evaluate participants' beliefs about their own and others' present bias. The experiment centers around a real-effort task, and the participants' predictions of their own and others' work decisions allow me to measure their beliefs about their own and others' present bias.

The experiment runs over the course of four weeks, recruiting participants from the Harvard Decision Sciences Lab. Each participant chooses a day of the week on which to participate, and needs to log in on that day of the week during each of the following four weeks. The instructions are presented on the first participation date, and the participants must pass a comprehension quiz in order to be eligible for the study. All instructions, questions, and assignments are catalogued in Appendix C; the informed consent language can be seen in Appendix D.

I present the experimental design in five subsections. First, I describe the experimental task and the information that the participants receive about other participants in the experiment. Next, I present the experimental timeline and detail the payment scheme. The work decisions faced by the participants are detailed in the third subsection, while the predictions are discussed in the fourth subsection. The last subsection presents information on the experimental sample, including sessions, recruitment, and attrition.

2.1 Experimental task

The real-effort task consists of a random sequence of characters appearing (one by one) on an otherwise empty screen, where participants are asked to press a key every time an asterisk appears. The duration of each round is 60 seconds: 50 seconds of work (with a total of 25 characters appearing during that period), followed by a 10 second break. The participants must achieve an average accuracy of 80% across all rounds within a session to successfully complete the work and receive payment. Figure 1 displays a sample task screen.

This task is specifically designed with a two-fold objective. First, the task needs to be tedious, so that the participants are exposed to the dynamic tension between the cost of completing more rounds of the task now and the benefit of receiving a higher payment later. Second, the task must be relatively straightforward and simple to complete, so that there is no skill involved, ensuring that any differences between predictions of the participants' own and others' choices are indeed driven by a wedge in beliefs about present bias, rather than overconfidence regarding skill.

While the character-identification task satisfies the objectives of being tedious and not requiring any skill, it is somewhat artificial, which poses a concern that participants might be ill-equipped to make predictions regarding either their own or others' behavior. In order to alleviate this concern and ensure that the elicited beliefs reflect real-world belief formation, I do the following:

1. All participants try the task for 5, 10, or 15 minutes before making any predictions, which ensures that they are familiar with what it is like to engage in the task.

2. A pilot study of the experimental design is run in October-November 2015, with participants recruited from the same Harvard Decision Sciences Lab pool as in the subsequent main experiment. Demographic data are gathered from all pilot study participants on the first participation date. Data on task enjoyment are gathered from the participants who complete the pilot study during a debrief questionnaire at the end of the last participation date.

3. Participants in the main experiment are presented with the data from the pilot study participants in an interactive display with break-downs by gender, race, marital status, age, education, and employment. A screenshot of this display is shown in Figure 2. The participants are encouraged to study these data as part of familiarizing themselves with the task.

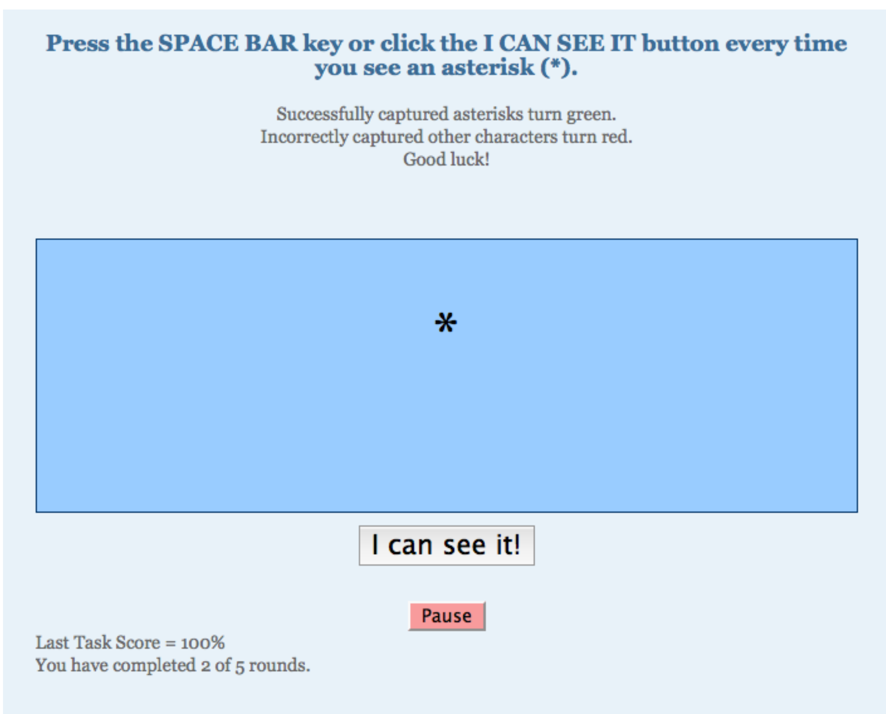


Figure 1: Screenshot of the experimental task.

Thus, when the participants are asked for predictions about others, they have some empirical familiarity with who the others are and how they feel about the task. The elicited beliefs then more closely correspond to beliefs in real-world scenarios, where individuals have familiarity with the general population of others and the assignment at hand.

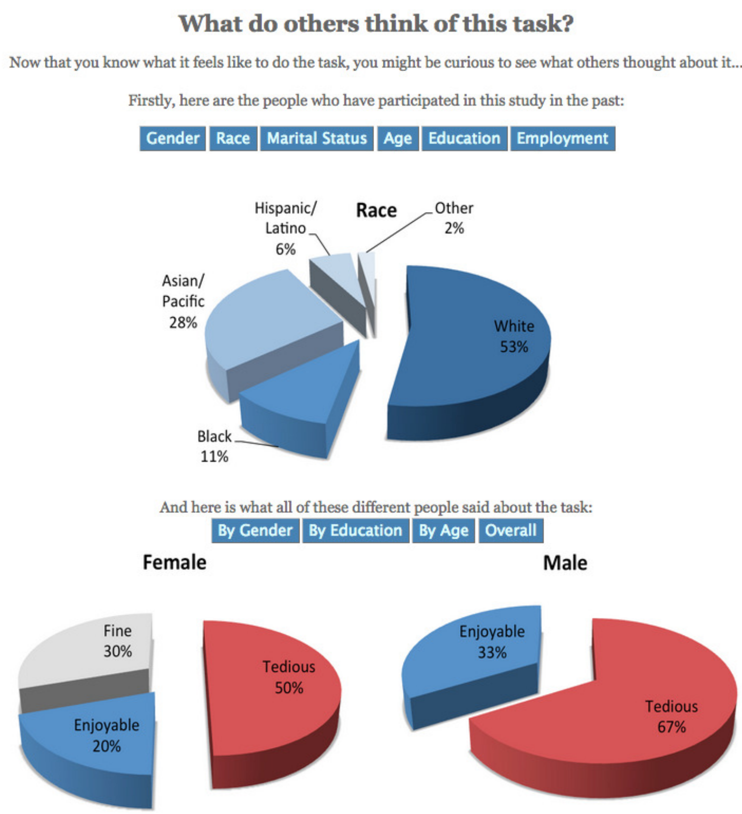


Figure 2: Breakdown of demographics and task enjoyment responses from the pilot study. This interactive display is shown to the participants in the online laboratory experiment to provide them with context regarding the backgrounds of other participants in the study and these other participants' opinions of the experimental task.

2.2 Experimental Timeline and Payments

Each participant logs into the experiment on her chosen day of the week during four consecutive weeks, denoted by Week 1, Week 2, Week 3, and Week 4. On each participation date, the participant must complete a mandatory warm-up of the task and answer all questions. At the end of the experiment, participants are paid based on the amount of work they do as well as completion of all mandatory items. The full experimental timeline is presented in Figure 3.

The first item on each participation date is a warm-up, which involves the participants having to do a mandatory number of rounds of the task. The warm-up amounts vary randomly across participants and consist of 5, 10, or 15 rounds. The differential warm-up amounts allow me to control for projection bias (see Loewenstein, O'Donoghue, and Rabin, 2003), which might lead participants to underestimate the effort cost of doing the task when not significantly exposed to it.

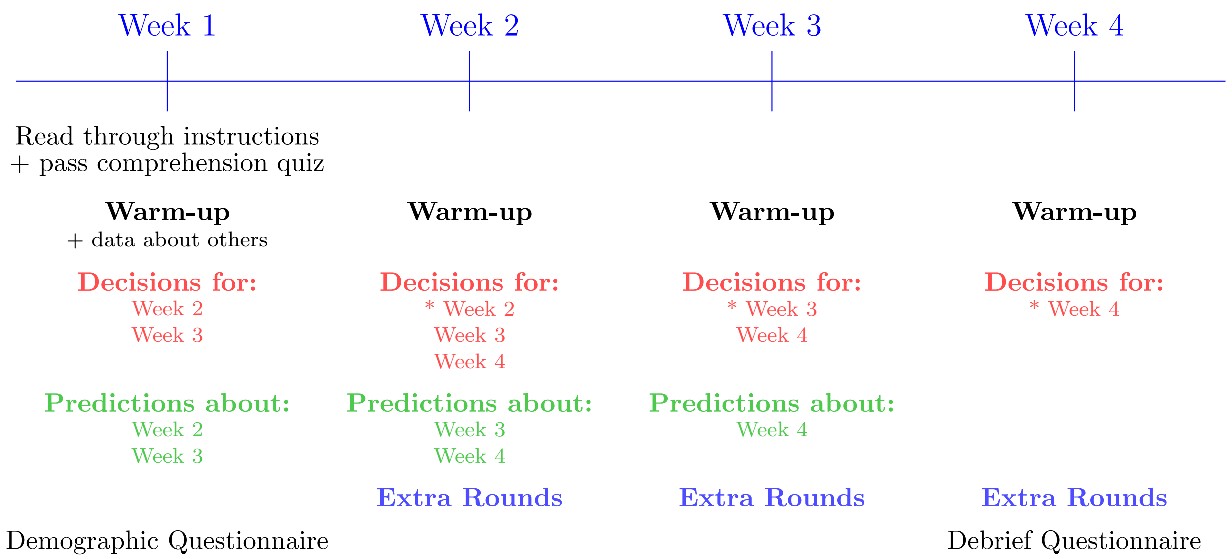


Figure 3: Experimental timeline. Stars indicate questions regarding decisions for immediate work.

After the warm-up, participants are asked how many extra rounds (between 0 and 70) of the task they would like to complete for additional pay at different wages, either on the same day or on future participation dates. The participants are also randomized into groups asked to predict either their own future decisions, or the average decisions of the other participants, or both.

Once all questions are answered, the next step is completing the chosen number of rounds of work. In particular, one of the decisions *for* the current date (made either *on* that date or earlier) is selected at random to be implemented, and the participant must immediately complete the number of extra rounds in that decision.

On the first participation date, participants also fill out a questionnaire consisting of demographic questions and questions eliciting the participants' predictions regarding their own and their peers' psychological state, time constraints, and preferences over the next few weeks, presented in random order. Similarly, at the end of the last participation date, there is a short debrief questionnaire. Participants are asked for reasons behind their own and others' inconsistencies, as well as predictions on whether they would behave more consistently if offered another chance to participate. The debrief also elicits beliefs regarding one's own and one's peers' present bias in other domains: expected gym attendance, work procrastination, and healthy eating.

Each participant's payment consists of two components: the completion payment and supplemental wages. The participant receives the $30 completion payment for logging in and completing all required work on each participation date. The supplemental wages are computed at the corresponding rates for any extra rounds that the participant completes, and any incentive bonuses earned for correct predictions. In order to be eligible for the $30 completion payment, the participant must complete each warm-up, answer all decision and prediction questions, and then finish the additional rounds in her implemented work decisions. If the participant fails to complete any of these tasks on one of her participation dates, the participant is disqualified and foregoes the $30 completion payment. Disqualified participants still receive payment for the additional rounds that they have completed before disqualification. The payments are dispensed in the form of Amazon.com gift cards on the Sunday one week after the end of Week 4.

2.3 Work Decisions

A critical component of the experimental design consists of the participants' decisions about how much of the real-effort task to do. The participants are asked to make these decisions for the current date and for future participation dates, and all decisions have an equal chance of being implemented. The differences in the participants' decisions for immediate versus future work are used to capture the participants' present bias.

The participants face work decisions on each of their four participation dates, indicated in red in Figure 3. Each set of decisions consists of two questions for the same date but at different wages. The wages are drawn randomly from between $0.10/round and $0.30/round, in increments of $0.05. This corresponds to $6/hour-$18/hour. All wages are equally likely to be drawn, with the restriction that the two wages on a single screen must be different. An example of a work decision screen is presented in Figure 4.

A few checks are in place to ensure that the participants' work decisions reflect their genuine preferences. First, if a participant enters the same number into both fields, she sees a warning enquiring whether she is certain that she would like to proceed with a decision to do the same amount of work regardless of the wage, or if she would like to reconsider. Second, if a participant enters a higher number into the field with the lower wage, she is asked whether she would really wish to do more work for lower pay, or whether she would like to reconsider her answers. These checks are in place to ensure that the participants are paying attention to the questions, rather than quickly entering random or repeating numbers into the fields. To ensure that the decisions correspond to permissible amounts, the participants must also enter an integer between 0 and 70 into each field to proceed.

Overall, each participant makes 16 work decisions – 6 immediate decisions for the same date and 10 ahead-of-time decisions for future dates. The full set of decisions is displayed in Panel 1 of Table 1, with each cell corresponding to a two-question screen analogous to the one displayed in Figure 4. The blue row catalogues immediate decisions, while the green rows list ahead-of-time decisions.

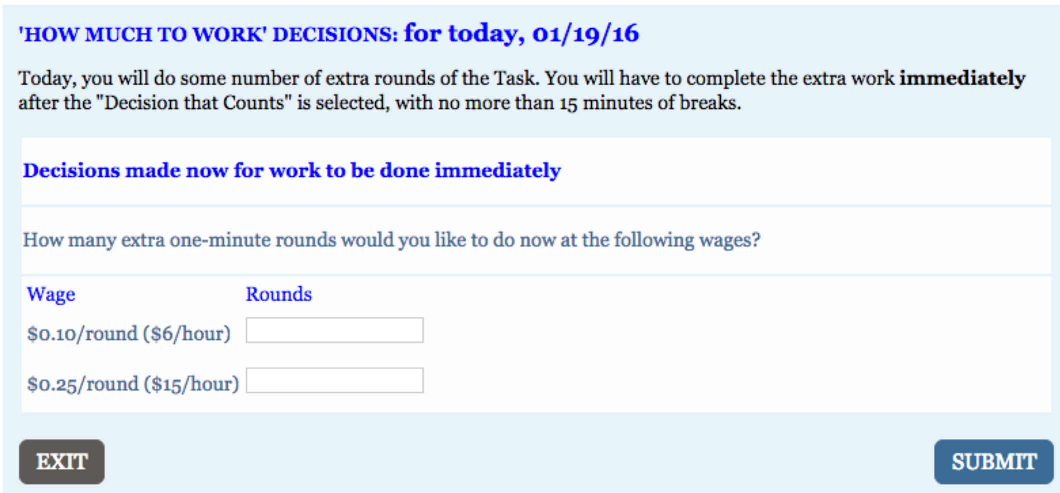


Figure 4: An example of a work decision screen.

During Weeks 2, 3, and 4, each participant's actual amount of work is randomly selected from all of the decisions that the participant has made for that date. In particular, once her work decisions are complete on a given participation date, the participant is shown all decisions that she has made for that date – this includes decisions made only moments earlier as well as ahead-of-time decisions made on prior dates. Figure 5 displays a sample screen aggregating all work decisions for a particular date.

The participant is reminded that, once one of these decisions is randomly selected to be implemented, she must complete the work in that decision immediately, with no more than a total of 15 minutes of break. This restriction serves to ensure that the decisions made for immediate work are perceived as truly immediate, rather than, for example, decisions made in the morning for work to be completed in the evening.

Once the participant clicks on the "SELECT" button, a randomization is run and one of the decisions is selected as the "Decision that Counts." All decisions are ex ante equally likely to be selected. The selected decision is then marked in dark blue, and a counter appears on the webpage. In order to continue participating in the experiment and receive the completion payment, the participant must complete the amount of work in the selected decision immediately with no more than 15 minutes of breaks.

2.4 Predictions

In order to compare participants' beliefs about their own and others' present bias, the participants are asked to make a set of predictions. After making their work decisions, the participants are asked to predict either how much work they will choose for immediate completion on future dates, or how much work other participants will choose, or both.

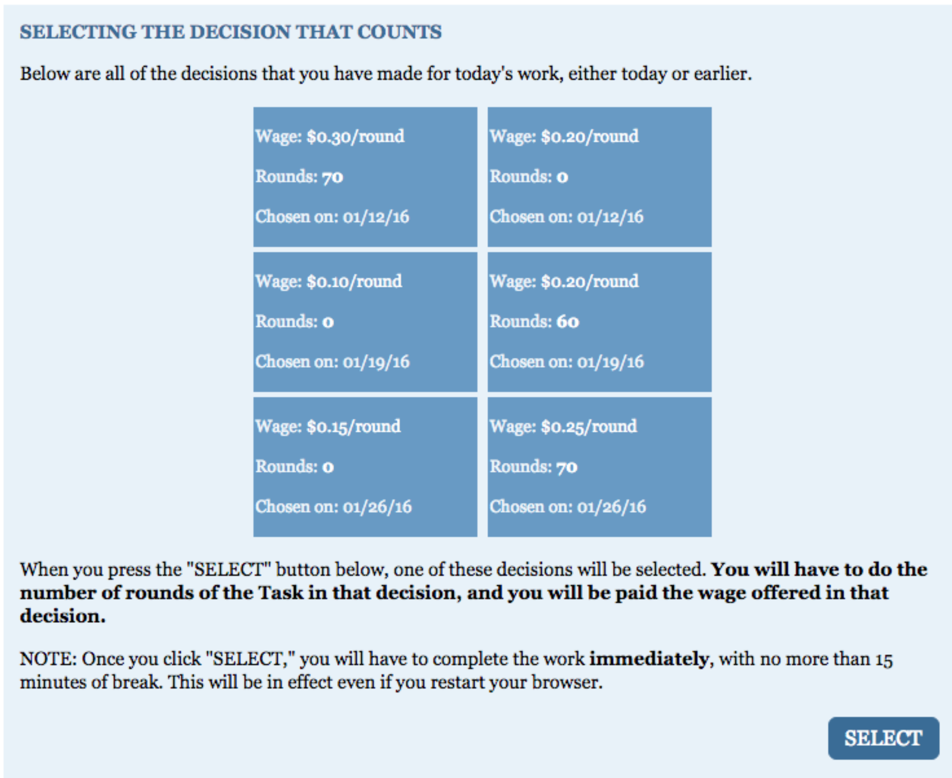


Figure 5: An example screen aggregating a participant's work decisions for a given date. Once the participant presses "Select," one choice is implemented as the "Decision that Counts." The participant then needs to complete the work from that decision *immediately,* with no more than 15 minutes of break.

For the prediction questions, I split the participants into the following three groups:

- **Group 1:** Throughout the experiment, participants in this group are asked how much work they anticipate choosing for immediate completion when various future dates actually arrive. Since the decision and prediction questions are quite similar, predictions appear side by side with the decision questions, in order to make the questions clearer and more straightforward. See Panel 1 of Figure 6 for an example screen presented to participants in this group.

- **Group 2:** Throughout the experiment, participants in this group are asked to predict how others make decisions. They are asked to predict the average of the other participants' current decisions for work on future dates, as well as the average of the other participants' choices for immediate completion when those future dates actually arrive. Participants in Group 2 are asked to make the predictions about others' current and future decisions side by side, as illustrated in Panel 2 of Figure 6.

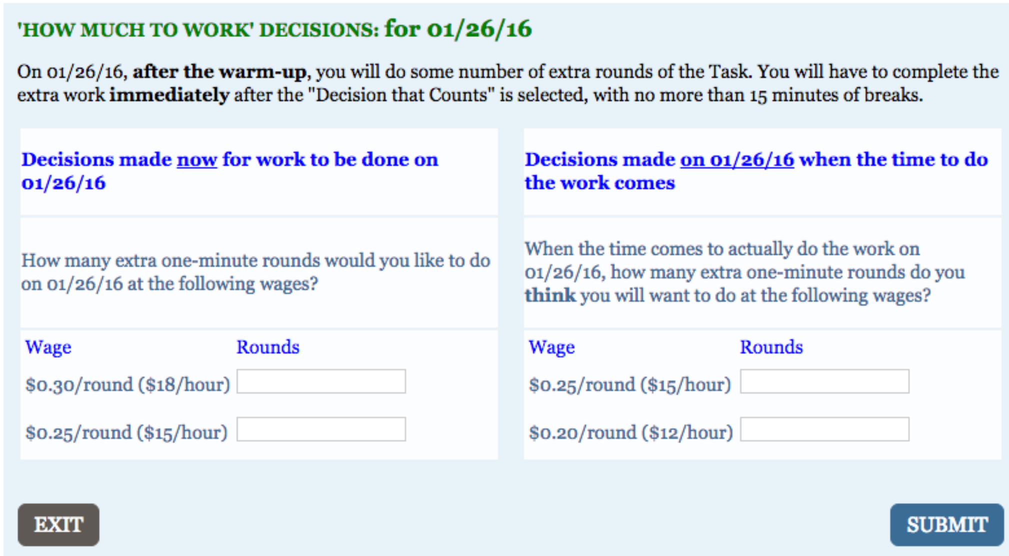
- **Group 3:** Participants in this group are asked both sets of questions described above and illustrated in Figure 6. The order in which the participants see these questions is randomized across participants.

I test the robustness of the participants' predictions to posing the two sets of questions (about self and others) to two separate groups of participants (Groups 1 and 2) versus to the same participants (Group 3). On the one hand, asking participants to make predictions about both themselves and others may lead to anchoring effects analogous to those documented by Tversky and Kahneman (1974), where participants use their answers to the first set of predictions as an anchor for the second set of predictions. In this sense, the answers by participants in Groups 1 and 2 present cleaner, unanchored beliefs regarding self and others. On the other hand, the juxtaposed answers of participants in Group 3 more accurately reflect beliefs in situations where individuals explicitly evaluate themselves and others in the same context. Such scenarios arise in a variety of common environments, including relative performance compensation contracts in the workplace and curve-graded assignments in schools. As I show in the next section, the effects are consistent across posing the two sets of questions separately and together, suggesting that the above concerns do not play a significant role for elicited beliefs.

The structure of the decision and prediction questions for both groups of participants is illustrated in Table 1. Present bias can be estimated by comparing immediate decisions (blue row in Panel 1 of Table 1) to ahead-of-time decisions (green rows in Panel 1). Beliefs about one's own present bias are captured by comparing one's ahead-of-time decisions for future dates (green rows in Panel 1) against predictions of one's decisions when the future dates actually arrive (blue row in Panel 2 and the first set of blue rows in Panel 4). Beliefs about others' present bias are estimated by comparing predictions of others' ahead-of-time decisions for future dates (green rows in Panels 3 and 4) against predictions of others' decisions when the future dates actually arrive (blue rows in Panel 3 and the second set of blue rows in Panel 4).

I wish to elicit thoughtful, truthful answers to the prediction questions. For predictions regarding others, this can be achieved by making the questions incentive-compatible with monetary rewards for correct predictions. Predictions about one's own behavior, however, are more subtle. In this case, there are feedback effects, since the correctness of these predictions is influenced by the participants' own subsequent behavior, which creates scope for strategic rather than truthful answers and behaviors. For example, participants may use their predictions as commitment devices to guide their future behavior.

Panel 1: Self-Predictions



Panel 2: Other-Predictions

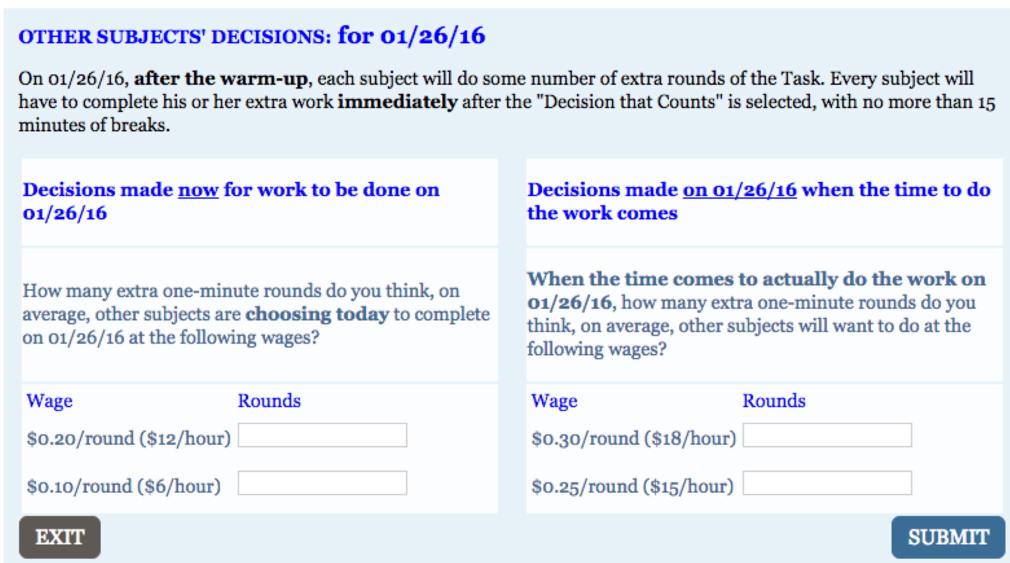


Figure 6: Examples of screens eliciting participants' predictions of their own and others' work decisions. **Panel 1** offers an example of a self-prediction screen. On the left, the participant is asked to make decisions for a future participation date. On the right, she is asked to predict how much work she will choose when that future participation date actually arrives. **Panel 2** presents an example of predictions about others. On the left, the participant is asked to guess what others are choosing now for the future. On the right, the participant is asked to predict how much work others will wish to do when the future date actually arrives.

Table 1: Decision and prediction questions posed to participants in Groups 1, 2, and 3. Decisions for immediate work and predictions of such decisions are marked in blue. Decisions for future work and predictions of such decisions are marked in green.

Panel 1: Decisions – All Participants

Decisions on Date 1	Decisions on Date 2	Decisions on Date 3	Decisions on Date 4
	For Date 2	For Date 3	For Date 4
For Date 2	For Date 3	For Date 4	
For Date 3	For Date 4		

Panel 2: Predictions – Group 1 Participants

Predictions on Date 1	Predictions on Date 2	Predictions on Date 3	Predictions on Date 4
Own decision on Date 2 for Date 2	Own decision on Date 3 for Date 3	Own decision on Date 4 for Date 4	
Own decision on Date 3 for Date 3	Own decision on Date 4 for Date 4		

Panel 3: Predictions – Group 2 Participants

Predictions on Date 1	Predictions on Date 2	Predictions on Date 3	Predictions on Date 4
Others' dec. on Date 2 for Date 2	Others' dec. on Date 3 for Date 3	Others' dec. on Date 4 for Date 4	
Others' dec. on Date 3 for Date 3	Others' dec. on Date 4 for Date 4		
Others' dec. on Date 1 for Date 2	Others' dec. on Date 1 for Date 3	Others' dec. on Date 1 for Date 4	
Others' dec. on Date 1 for Date 3	Others' dec. on Date 1 for Date 4		

Panel 4: Predictions – Group 3 Participants

Predictions on Date 1	Predictions on Date 2	Predictions on Date 3	Predictions on Date 4
Own decision on Date 2 for Date 2	Own decision on Date 3 for Date 3	Own decision on Date 4 for Date 4	
Own decision on Date 3 for Date 3	Own decision on Date 4 for Date 4		
Others' dec. on Date 2 for Date 2	Others' dec. on Date 3 for Date 3	Others' dec. on Date 4 for Date 4	
Others' dec. on Date 3 for Date 3	Others' dec. on Date 4 for Date 4		
Others' dec. on Date 1 for Date 2	Others' dec. on Date 1 for Date 3	Others' dec. on Date 1 for Date 4	
Others' dec. on Date 1 for Date 3	Others' dec. on Date 1 for Date 4		

To check that the monetary incentives do not prompt any commitment demand that would perversely affect participants' self-predictions, I randomly assign each participant into either the incentivized or the unincentivized treatment arm, with equal probability. Participants in the incentivized arm are given a monetary incentive for correct predictions about decisions that are eventually implemented. The monetary incentives are randomized across these participants, and vary from \$0.10 to \$0.40 – similar to the wages for one minute of work. Participants in the unincentivized arm are asked to state their predictions without any monetary incentive. In order to keep the design symmetric, this is implemented analogously for participants making predictions about self, those making predictions about others, and those making both sets of predictions. The incentive structure extends equally to all predictions made by a given participant.

For example, consider a participant from Group 1 or 3 who is randomly assigned to the incentivized group with a prediction bonus of \$0.20. Suppose that she is asked on her first participation date (Date 1) to predict how much work she will choose to do immediately at \$0.10/round on Date 2, and she answers 15 rounds. Then she receives a prediction bonus of \$0.20 if the following conditions are met: (a) on Date 2, she is asked how much work she would like to complete immediately at \$0.10/round, and she chooses 15 rounds; and (b) this decision is implemented as the "Decision that Counts." Similarly, consider a participant from Group 2 or 3, who is randomly assigned to the incentivized group with a prediction bonus of \$0.20. Suppose that on Date 1, she is asked to predict how much work, on average, other participants will prefer to do immediately on Date 2 at \$0.10/round, and she answers 15 rounds. Then she will receive a bonus of \$0.20 if: (a) on Date 2, at least one other participant is asked how much work he would like to complete immediately at \$0.10/round, and the average answer is 15 rounds; and (b) this decision is implemented as the "Decision that Counts" for at least one of the other participants.

2.5 Sample

The experiment runs over five non-overlapping sessions, with a total of 364 individuals taking part, recruited through the Harvard Decision Sciences Lab. In order to be eligible for the study, participants must pass a comprehension quiz after reading the instructions, testing the participants' understanding of the experiment. The comprehension quiz includes questions regarding payment, timeline, decisions, and predictions; the full quiz is catalogued in Appendix C.

The goal of running the experiment over multiple sessions is to minimize the effects of any unforeseen systematic shocks such as weather disruptions or university-wide events. Furthermore, since a large part of the Harvard Decision Sciences Lab subject pool consists of Harvard University undergraduates, the five sessions are explicitly timed to avoid the University's midterm exams (mid-March) and final exams (May). The five experimental sessions are run at the following times:

- Session 1: January 11 - February 7, 2016

- Session 2: February 8 - March 6, 2016
- Session 3: March 28 - April 24, 2016
- Session 4: June 6 - July 3, 2016
- Session 5: July 11 - August 7, 2016

In addition, a small-scale pilot study of the experimental design is run during October 12 - November 8, 2015. For the results of the pilot study, please refer to Appendix A.2.

Table 2: Numbers of recruited participants and attrition rates across experimental sessions.

	Consent	Week 1	Week 2	Week 3	Week 4
Pilot	27	23	21	19	19
Sessions 1-5	364	278	230	208	198
Session 1: Jan. 11 - Feb. 7, 2016	78	61	53	50	50
Session 2: Feb. 8 - Mar. 6, 2016	81	65	59	52	50
Session 3: Mar. 26 - Apr. 24, 2016	86	71	57	49	43
Session 4: Jun. 6 - Jul. 3, 2016	64	42	31	29	28
Session 5: Jul. 11 - Aug. 7, 2016	55	39	30	28	27

A total of 198 participants complete the entirety of the experiment during the five experimental sessions, with an additional 166 participants consenting to participate but not finishing the entirety of the four-week-long experiment. A break-down of recruited participants and attrition rates by session is reported in Table 2. Since registering for the online study is virtually costless, a large number of participants drop out once they begin reading the instructions upon their first log in; of the 364 participants consenting to take part in the experiment, 86 (24%) do not complete the instructions, warm-up, and work decisions on the first participation date or fail the comprehension quiz. The attrition rates attenuate over the subsequent weeks, as exiting the experiment costs the participants their \$30 completion payments. Of the 278 participants who finish their first participation date, 230 (83%) complete the second participation date, 208 (75%) complete the third participation date, and 198 (71%) complete the entirety of the experiment. The results detailed below are robust to including attrited participants and to focusing solely on those participants who complete the entirety of the experiment.

3 Reduced-Form Results

The reduced-form results from the laboratory experiment suggest that participants have more awareness regarding present bias in others than in themselves. The results are robust to varying the incentive structure for the predictions and to posing the questions about self and others together to the same participants or to separate groups of participants.

3.1 Present Bias and Beliefs

Pooled results from all participants who finish the experiment indicate that (i) participants display present bias in their effort choices; (ii) participants do not anticipate their own present bias; and (iii) participants expect present bias in others. For robustness of these results to the inclusion of attrited participants who complete the preliminaries but do not finish the experiment, please refer to Appendix A.1.

The pooled sample consists of the 198 participants who complete the entirety of the experiment. Of these, 60 are in Group 1, 60 are in Group 2, and 78 are in Group 3. Each participant is asked to make a total of 6 decisions for immediate work and 10 decisions for future dates. In addition, each participant in Groups 1 and 3 answers 10 questions regarding her own decisions when the future dates actually arrive. Similarly, each participant in Groups 2 and 3 is asked a total of 10 questions regarding others' current work decisions for future dates and 10 questions about what others will choose when the future dates actually arrive.

Present bias is estimated by comparing participants' work decisions for future dates against their work decisions for immediate completion. The experimental participants choose to do, on average, 30.03 rounds per session when the choices are elicited ahead of time. When asked how much work they would like to complete immediately, the participants choose an average of 26.53 rounds per session. Figure 7 plots the ahead-of-time and immediate work decisions across the five possible wages from \$0.10/minute to \$0.30/minute, with standard error bars clustered by participant. As illustrated in the figure, participants choose to do more work when the decision is made in advance for all wages except for \$0.10/minute.

The difference between the two types of decisions is statistically significant and robust to controlling for wage fixed effects and participant fixed effects. The results are presented in Panel 1 of Table 3. Participants choose to do, on average, 3.50 rounds fewer when their decisions are for immediate work (3.33 rounds when controlling for wage fixed effects, 3.50 with participant fixed effects, and 3.36 including both fixed effects). The difference is statistically significant at the 1% level across specifications, and consistent with prior evidence on present bias in real-effort tasks (see, e.g., Augenblick et al., 2015).

Are the participants aware of this time inconsistency in their effort choices? The participants' naïveté regarding their own present bias is captured by comparing their work decisions for future dates against their *predictions* of the decisions they will make when those dates actually arrive. Forecasts of lower work decisions when the dates actually arrive would indicate experimental participants' sophistication regarding their present bias. On the other hand, if participants do not anticipate their decisions changing when the work becomes imminent, then they display naïveté regarding their present bias.

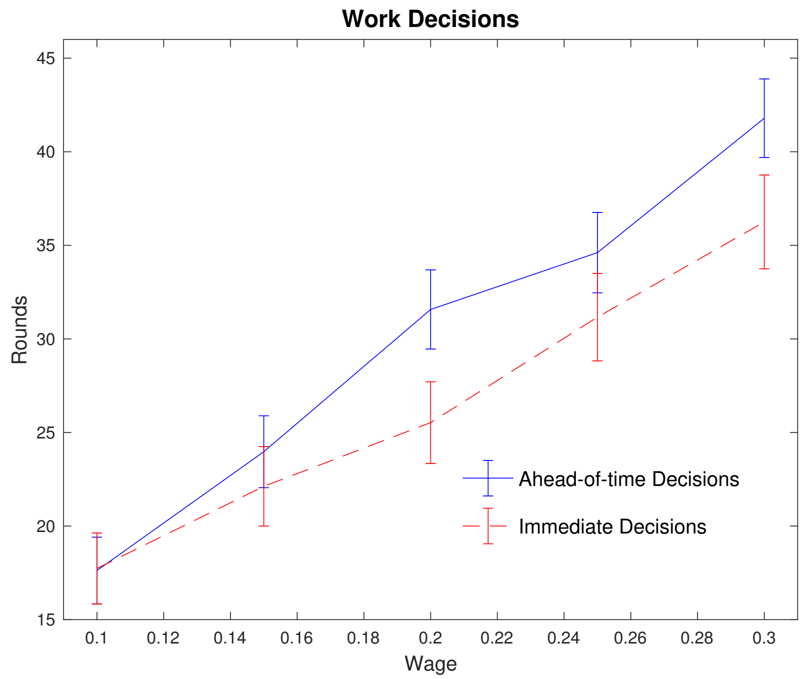


Figure 7: A comparison of work decisions for the current date (immediate decisions) and for future dates (ahead-of-time decisions). The decisions are displayed for each of the five possible wages between \$0.10/round and \$0.30/round.

The results indicate that, on average, experimental participants display little anticipation of their own present bias. The average predicted differences in their effort choices, estimated from the 138 participants making self-predictions (i.e., participants from Groups 1 and 3), are displayed in Panel 2 of Table 3. The differences are small, and significant only in one specification (participant fixed effects, no wage fixed effects). The differences also do not have consistent signs across the specifications. This is consistent with the findings of naïveté in Augenblick and Rabin (2018).

By contrast, predictions about others, made by the 138 participants in Groups 2 and 3, reveal awareness of others' present bias. The experimental participants expect their peers to choose less work when the decisions are for immediate completion than when the decisions are for future completion. Asked how many rounds others wish to do ahead of time, participants predict an average of 28.28 rounds. When asked about others' work decisions for immediate completion, the average prediction is 27.01 rounds. The difference, estimated in Panel 3 of Table 3, is significant at the 1% level if participant fixed effects are included, and at the

5% level omitting participant fixed effects. The results are robust to inclusion of attrited participants, and are statistically stronger in this larger sample (see Appendix A.1).

Table 3: Reduced-form results from participants who complete the entirety of the experiment. **Panel 1** estimates the actual difference between work decisions made ahead of time and work decisions for immediate completion. **Panel 2** estimates the predicted difference in ahead-of-time vs. immediate decisions when participants are asked to make predictions about themselves. **Panel 3** displays the predicted difference in ahead-of-time vs. immediate decisions when participants are asked to make predictions for other experimental participants. The differences are estimated with and without wage and participant fixed effects, and standard errors are clustered by participant.

Panel 1: Actual Difference in Ahead-of-time vs. Immediate Decisions				
Actual Difference	**3.50**	**3.33**	**3.50**	**3.36**
Standard error	*(0.59)*	*(0.55)*	*(0.61)*	*(0.57)*
Controls:				
Wage FE		X		X
Participant FE			X	X

Panel 2: Predicted Difference in Own Ahead-of-time vs. Immediate Decisions				
Self-Prediction	**0.10**	**-0.43**	**1.15**	**0.62**
Standard error	*(1.13)*	*(1.12)*	*(0.40)*	*(0.34)*
Controls:				
Wage FE		X		X
Participant FE			X	X

Panel 3: Predicted Difference in Others' Ahead-of-time vs. Immediate Decisions				
Other-Prediction	**1.27**	**1.47**	**1.27**	**1.49**
Standard error	*(0.52)*	*(0.42)*	*(0.53)*	*(0.44)*
Controls:				
Wage FE		X		X
Participant FE			X	X

Interestingly, while predictions regarding others reveal that participants expect present bias in others, participants do not correctly guess the magnitude of the effect. The average predicted difference in others' decisions is 1.49 rounds, whereas the actual average difference in participants' work decisions is 3.50 rounds. This contrasts with the classroom experiment detailed in Section 5, in which the students' predictions regarding their peers are remarkably well calibrated. The difference in accuracy across the two settings is most likely attributable to the participants' different levels of experience with the two settings: students have substantial experience observing their classmates procrastinate on assignments, but participants in the experiment have no experience observing others choose work decisions for the experimental task.

Overall, the reduced-form results from the online experiment paint the following picture. Experimental participants display almost full naïveté about their own present bias but some, although imperfect, awareness of others' present bias. In the remainder of this section, I explore the robustness of these findings to incentivizing predictions and to posing the questions jointly or separately.

3.2 Incentivizing Predictions

I test the robustness of the results to altering the incentivization mechanism for eliciting participants' predictions regarding their own and others' work decisions. The predictions are not significantly different when the questions are posed in an unincentivized manner versus when the participants are offered monetary bonuses for correct predictions.

The incentive structure is randomized across experimental participants. Each participant is randomly allocated, with equal probability, to either the incentivized or the unincentivized treatment arm. Within the incentivized arm, the size of the incentive is randomly selected from \$0.10, \$0.20, \$0.30, or \$0.40 per correct prediction, with equal likelihoods. Thus, of the 198 participants who finish the experiment, 95 are unincentivized, and 103 are incentivized, with 23 participants receiving the \$0.10 bonus, 24 receiving the \$0.20 bonus, 30 receiving the \$0.30 bonus, and 26 receiving the \$0.40 bonus.

The participants do not display significant awareness of their own present bias, regardless of the incentive structure. Panel 1 of Table 4 reports the average predicted differences in one's own decisions for the 72 participants asked to make incentivized self-predictions and the 66 participants making self-predictions without monetary incentives. The predicted differences are estimated with wage and participant fixed effects. The average predicted difference is 0.59 rounds with the incentive and 0.65 rounds without the incentive. In both cases, the predicted difference is statistically indistinguishable from zero, indicating that the participants do not anticipate present bias in their own decisions, regardless of whether they are incentivized for correct predictions. For the incentivized group, the participants' naïveté is likewise robust across the size of the incentive. With the exception of the \$0.10 incentive, for which the predicted difference is 0.03 rounds, the point estimates of the predicted differences are approximately 0.60 rounds across the incentive amounts. For none of the incentives are these predicted differences statistically distinguishable from zero, although the sliced samples are too small to properly evaluate significance.

Incentivizing predictions also has no significant effect on the elicited beliefs about other participants. The predicted differences in others' decisions are, on average, 1.41 rounds when the predictions are incentivized and 1.70 rounds without the incentive. In both cases, the predicted differences are statistically different from zero. The former is significant at the 5% level and the latter at the 1% level.

Overall, the subsample analysis slicing by incentive indicates that the results are not driven by strategic responses to incentive structures. Instead, the participants' answers are robust to incentivized and unincentivized elicitation of beliefs. Across the board, participants display fairly precise awareness of others' present bias, and no significant awareness of their own present bias.

Table 4: Predicted differences in one's own and others' decisions, sliced by incentive. **Panel 1** evaluates self-predictions for participants who receive monetary incentives for correct predictions and those who do not. **Panel 2** further slices self-predictions of incentivized participants by the size of the incentive. **Panel 3** analyzes predictions regarding others, when responses are elicited with and without monetary incentives. The differences are estimated with wage and participant fixed effects, and standard errors are clustered by participant.

Panel 1: Self-Predictions by Incentivized and Unincentivized Participants

	Incentivized	Unincentivized
Self-Prediction	**0.59**	**0.65**
Standard error	*(0.48)*	*(0.46)*
Controls:		
Wage FE	X	X
Participant FE	X	X

Panel 2: Self-Predictions by Incentivized Participants, Varying Size of Incentive

	\$0.10 incentive	\$0.20 incentive	\$0.30 incentive	\$0.40 incentive
Self-Prediction	**0.03**	**0.66**	**0.67**	**0.57**
Standard error	*(0.98)*	*(1.11)*	*(0.58)*	*(0.98)*
Controls:				
Wage FE	X	X	X	X
Participant FE	X	X	X	X

Panel 3: Other-Predictions by Incentivized and Unincentivized Participants

	Incentivized	Unincentivized
Other-Prediction	**1.41**	**1.70**
Standard error	*(0.62)*	*(0.60)*
Controls:		
Wage FE	X	X
Participant FE	X	X

3.3 Juxtaposing Predictions about Self and Others

Next, I confirm the robustness of the results to posing the two sets of questions (predictions regarding self and others) to the same participants or separately to two groups of participants. Participants' answers do not systematically vary across the two methods of posing the questions, as evidenced by the results in Table 5.

Participants do not expect significant present bias in themselves, regardless of whether they are also asked to make predictions about others. The 60 participants who make only self-predictions anticipate that they will choose to do an average of 0.62 rounds fewer when the work decision has immediate consequences; this predicted difference is not statistically different from zero. Similarly, the 78 participants who also face questions about others predict that they will choose an average of 0.73 rounds fewer for immediate completion, also statistically indistinguishable from zero.

Likewise, participants' expectations of others' present bias are similar across those who only make predictions regarding others and those who answer both sets of questions. The 60 participants who are only asked to make predictions about others expect that the average other participant will want to do 1.53 rounds fewer when the work decision is made for immediate completion. The 78 participants who are asked to make both sets of predictions anticipate that others will choose 1.46 rounds fewer when the work decision is made for immediate completion. The predicted differences by both groups are statistically significantly different from zero at the 5% level.

Overall, participants appear to be providing independent answers for the two sets of questions, and their answers are robust to seeing only one type of question or both. These results suggest that the participants' predictions regarding their own future decisions and regarding the decisions of the other participants are not influenced by anchoring effects or strategic comparisons. Instead, the elicited predictions reflect the participants' underlying beliefs regarding self and others.

In Appendix A.3, I exploit differences in participants' warm-up amounts to explore whether the warm-up amount systematically affects the participants' displayed present bias and predictions, and find mixed results. Overall, the analysis sliced by warm-up indicates two patterns. First, my experimental design does not elicit significant projection bias in the participants. Second, the qualitative patterns of more critical beliefs regarding others tend to hold across the different warm-up amounts.

Table 5: Predicted differences in one's own and others' decisions, sliced by method of presenting the questions. **Panel 1** considers participants' predictions of the differences in their own work decisions, elicited from participants who only make predictions about themselves (left column) and from participants who are asked both self- and other- prediction questions (right column). **Panel 2** presents predictions for others, elicited from participants who only make predictions regarding others (left column) and from participants who are asked both sets of questions (right column). The differences are estimated with wage and participant fixed effects, and standard errors are clustered by participant.

Panel 1: Self-Predictions		
	Self-Prediction Only	Both Sets of Questions
Self-Prediction	**0.62**	**0.73**
Standard error	*(0.41)*	*(0.50)*
Controls:		
Wage FE	X	X
Participant FE	X	X

Panel 2: Other-Predictions		
	Other-Prediction Only	Both Sets of Questions
Other-Prediction	**1.53**	**1.46**
Standard error	*(0.59)*	*(0.64)*
Controls:		
Wage FE	X	X
Participant FE	X	X

4 Structural Estimation

In this section, I present a model of decision-making and predictions within the framework of β-δ preferences and use my experimental data to estimate the model's parameters. I find estimates of present bias consistent with prior literature, virtually no awareness of one's own present bias, and robust albeit incomplete awareness of others' present bias.

4.1 Model

I begin by outlining the functional form assumptions for my structural estimation. I then present the resulting expressions for the participants' work decisions, their predictions regarding their own future decisions, and their predictions regarding the decisions of other participants.

Structural Assumptions. The model of individual decision-making and predictions is based on standard quasi-hyperbolic discounting. Each subsequent period is discounted relative to the preceding period by a time-consistent (daily) discount factor δ; in addition, all periods outside of the current one are discounted by the present bias parameter β.

I denote each participant's beliefs regarding her own present bias by $\beta_{(s)}$, and beliefs regarding others by $\beta_{(o)}$. Aggregate parameter estimation constrains the discount factors, beliefs, and all functional form parameters to be the same across all experimental participants.

I assume that each participant's utility function is separable in effort and money. Thus, when a participant performs e rounds of the task in addition to her x warm-up rounds, she incurs some effort cost $c(e+x)$ immediately and receives utility $U(w \times e)$ from the monetary payoff $w \times e$ on the predetermined future (post-experiment) payment date. For the purposes of estimation, I follow Augenblick and Rabin (2018) and assume a linear form for the utility in money with parameter ϕ (i.e., $U(w \times e) = \phi e w$) and a power form for the cost of effort with parameter γ (i.e., $c(e) = \frac{1}{\gamma} e^{\gamma}$).

When evaluating predictions regarding others, I assume that all participants perceive others' utility function to have the same functional form and parameters ϕ and γ as their own.[11] I also assume that a participant with a warm-up amount of x understands all other experimental participants to have the same warm-up x.[12] Thus, the only feature that the participants perceive as different for others than for themselves is the feature of interest: present bias.

Decisions. The structural assumptions allow for the following expression of the participants' decisions. Let T denote the payment date, and consider an individual with a warm-up amount x making a decision about how much work to complete immediately at wage w on date τ. She discounts the future payment by $\beta\delta^{T-\tau}$, but does not discount the immediate cost of effort. As a result, her chosen number of extra rounds is given by:

$$e^*_{imm} = \arg\max_e \{\beta\delta^{T-\tau}\phi w e - \frac{1}{\gamma}(e+x)^{\gamma}\} = (\beta\delta^{T-\tau}\phi w)^{\frac{1}{\gamma-1}} - x \tag{1}$$

Now suppose that the participant is again considering how much to work on the same date τ, but the decision itself is now made ahead of time, on date $t < \tau$. In this case, both the monetary payment and the effort cost are incurred in the future, and both are discounted by the present bias parameter β. As a result, the participant chooses the following number

[11] I relax this assumption in Appendix B.2, where I estimate self-predictions and other-predictions separately, allowing for different beliefs regarding others' parameters γ and ϕ. In Appendix B.1, I do not allow for differing γ and ϕ, but incorporate a difference in baseline levels of actual ahead-of-time work decisions and the predicted ahead-of-time decisions by others.

[12] Since the participants are not told that the warm-up amounts can vary, and there is no communication between experimental participants, there is no reason for them to think that others have different warm-up amounts.

of extra rounds:

$$e^*_{del} = \arg\max_e \{\beta\delta^{T-t}\phi we - \beta\delta^{\tau-t}\frac{1}{\gamma}(e+x)^\gamma\} = (\delta^{T-\tau}\phi w)^{\frac{1}{\gamma-1}} - x \tag{2}$$

Self-Predictions. In order to model a participant's expectations regarding her own future decisions, consider a participant on date t asked what choice she would make for immediate work at a wage w when a future date τ actually arrives. Effectively, the participant is predicting, ahead of time on date t, the decision expressed in (1).

Under the structural assumptions outlined above, the participant is aware of her effort cost function and her utility from money, but may hold incorrect beliefs regarding her present bias parameter β. In particular, the participant thinks that her future self will be making the decision in (1) under a present bias parameter $\beta_{(s)}$. Hence, she expects to choose the following number of extra rounds:

$$e^{(s)}_{imm} = \arg\max_e \{\beta_{(s)}\delta^{T-\tau}\phi we - \frac{1}{\gamma}(e+x)^\gamma\} = (\beta_{(s)}\delta^{T-\tau}\phi w)^{\frac{1}{\gamma-1}} - x \tag{3}$$

Other-Predictions. I now turn to expectations regarding the decisions made by other participants.

First, consider a participant in period t asked to predict how many rounds of the task others are choosing now for some future date τ, at a given wage w. The participant assumes that others have the same warm-up amount as her, x, as well as the same utility in money, effort cost function, and time-consistent discount factor δ. She perceives the others' present bias parameter to be $\beta_{(o)}$. Since the predicted choice is made for future work, the participant understands that others will discount both the effort cost and the monetary payoff by the present bias parameter. She hence perceives the others' decision for the delayed work as follows:

$$e^{(o)}_{del} = \arg\max_e \{\beta_{(o)}\delta^{T-t}\phi we - \beta_{(o)}\delta^{\tau-t}\frac{1}{\gamma}(e+x)^\gamma\} = (\delta^{T-\tau}\phi w)^{\frac{1}{\gamma-1}} - x \tag{4}$$

Second, consider a participant in period t predicting how many rounds of the task others will want to do when the future date τ actually arrives. Since the predicted choice consists of a trade-off between immediate work and a delayed monetary payoff (in period T), the participant expects that only the monetary payoff will be discounted by the present bias parameter $\beta_{(o)}$. As a result, she expects others to make the following choice:

$$e^{(o)}_{imm} = \arg\max_e \{\beta_{(o)}\delta^{T-\tau}\phi we - \frac{1}{\gamma}(e+x)^\gamma\} = (\beta_{(o)}\delta^{T-\tau}\phi w)^{\frac{1}{\gamma-1}} - x \tag{5}$$

The discussion above makes one key implicit assumption: that participants' predictions reflect their true expectations regarding their own and others' behavior. For predictions regarding others, this assumption should not pose any problems: participants have no strategic incentive to misstate their beliefs regarding others, and the prediction bonuses offer an unambiguous incentive to provide their best guesses. For predictions regarding one's own future behavior, the bonus might introduce a strategic incentive to use the predictions as a commitment device. However, the analysis in Section 4.2 indicates that the participants' answers do not vary significantly based on incentivizing predictions. Furthermore, Augenblick and Rabin (2018) find that allowing for the possibility that participants are optimally using their predictions as commitment devices does not alter the structural estimates of β and $\beta_{(s)}$. Hence, throughout this section, I assume that all predictions are stated truthfully.

4.2 Empirical Estimation

I now estimate the parameters in equations (1)-(5) using data from the experimental participants' responses to the decision and prediction questions.

Let $e(t, \tau, w, x, \mathbb{1}_s, \mathbb{1}_o)$ denote the stated number of extra rounds in response to a question posed to a participant with warm-up amount x on date t regarding work on date τ at wage w. The indicator variable $\mathbb{1}_s$ captures self-prediction responses, while the indicator variable $\mathbb{1}_o$ denotes responses about others. The remaining responses are the participant's actual decisions. The indicator $\mathbb{1}_{t=\tau}$ is equal to one if and only if the date when the response is elicited, t, is the same as the date for which the question is posed, τ.

Combining (1)-(5), I express the model's prediction of the participant's response, $\hat{e}(t, \tau, w, x, \mathbb{1}_s, \mathbb{1}_o)$, in terms of the model parameters:

$$\hat{e}(t, \tau, w, x, \mathbb{1}_s, \mathbb{1}_o) = (\beta^{\mathbb{1}_{t=\tau}} \beta_{(s)}^{\mathbb{1}_s} \beta_{(o)}^{\mathbb{1}_{t=\tau}\mathbb{1}_o} \delta^{T-\tau} \phi w)^{\frac{1}{\gamma-1}} - x \tag{6}$$

I estimate the parameters in (6) using maximum likelihood. Since the participants cannot pick fewer than 0 rounds or more than 70 rounds in any of the questions, I use a two-limit Tobit regression, with censoring from below at 0 and from above at 70 rounds.

Intuitively, the model parameters are identified from the data as follows:

- Variation in the timing of the decisions identifies β. In equation (6), the parameter β is present for immediate decisions ($\mathbb{1}_{t=\tau} = 1$ when $t = \tau$) but not for future decisions. Thus, present bias is identified by the difference between the number of rounds chosen for future participation dates and the number of rounds chosen for immediate work. Higher β corresponds to a smaller difference between these two sets of decisions.

- Restricting attention to the decisions made for the future, variation in the proximity of these future dates identifies the time-consistent discount factor δ, which enters multiplicatively with each day separating the decision date t from the work date τ.

- The parameter $\beta_{(s)}$ is identified through a comparison of the participants' actual decisions for immediate work and their predictions regarding these decisions, since the term $\beta_{(s)}$ enters only for the latter set of responses. Higher $\beta_{(s)}$ thus corresponds to smaller predicted differences in one's own choices.

- Similarly, the parameter $\beta_{(o)}$ is identified through a comparison of choices for future dates against participants' predictions regarding others' choices when those dates actually arrive, since only the latter set of responses reflects the parameter $\beta_{(o)}$.[13] The higher is $\beta_{(o)}$, the less difference participants anticipate to see in others' choices.

- Lastly, parameters γ and ϕ, which capture the shape of the participants' utility function in money and effort, are identified through variation in the wage w, which traces out the curvature and intercept of the utility function.

The parameter estimates are presented in Table 6. The first column reports the estimates of the two-limit Tobit regression of (6) restricting attention to the 198 participants who complete the entirety of the experiment. The second column includes all participants who provide at least one work decision or prediction response. Standard errors are bootstrapped and clustered by participant across the specifications. The estimated cost of effort is close to quadratic, with the estimate of $\hat{\gamma}$ at approximately 2.25, and the estimate of ϕ is around 380. The time-consistent discount factor δ is slightly above 1, indicating that when making decisions for the future, participants choose to do more rounds of work when the future is relatively nearer. This is possibly driven by participants being more certain regarding their schedules for the near future than for dates further in the future, and is consistent with the pattern documented by Augenblick and Rabin (2018).

The estimated model parameters reveal a significant extent of present bias among experimental participants. Estimates of the parameter β are around 0.82-0.86, with the difference from the null of $\beta = 1$ (no present bias) statistically significant at the 1% level. The estimates are consistent with prior evidence on present bias: for example, Laibson et al. (2008) estimate β around 0.71 using consumption choices, while Augenblick et al. (2015) document β around 0.89 for real-effort tasks. In the closest setting to the present paper, Augenblick and Rabin (2018) obtain estimates of β around 0.83.

[13] Note that in (6), the participants' current choices for future work and their predictions of others' current choices for future work are indistinguishable. I allow for the possibility of a difference in these baseline levels in Appendix B.1.

Table 6: Parameter estimates from the structural model. The table displays the estimates of the present bias parameter β, awareness of one's own present bias $\beta_{(s)}$, and beliefs regarding others' present bias $\beta_{(o)}$, as well as time-consistent discount factor δ, power effort cost parameter γ, and utility in money ϕ. I estimate the following model for predicted response $\hat{e}(t, \tau, w, x, \mathbb{1}_s, \mathbb{1}_o)$:
$\hat{e}(t, \tau, w, x, \mathbb{1}_s, \mathbb{1}_o) = (\beta^{\mathbb{1}_{t=\tau}} \beta_{(s)}^{\mathbb{1}_s} \beta_{(o)}^{\mathbb{1}_{t=\tau} \mathbb{1}_o} \delta^{T-\tau} \phi w)^{\frac{1}{\gamma-1}} - x$
The first column restricts attention to responses from participants who complete the entirety of the experiment. The second column includes responses from attrited participants. The estimation is done using a two-limit Tobit regression with censoring from below at 0 and from above at 70 rounds. Bootstrapped standard errors are clustered by participant in all specifications.

Parameter	Without attrited participants	With attrited participants
Present bias β	**0.8589**	**0.8151**
	(0.0330)	*(0.0335)*
Self-prediction $\beta_{(s)}$	**1.0502**	**1.0306**
	(0.0629)	*(0.0523)*
Other-prediction $\beta_{(o)}$	**0.8711**	**0.8715**
	(0.0349)	*(0.0314)*
δ	1.0147	1.0154
	(0.0024)	*(0.0028)*
γ	2.2485	2.2481
	(0.1123)	*(0.0910)*
ϕ	377.4181	385.9745
	(209.5252)	*(169.9324)*

Participants' beliefs regarding their own present bias, captured by the parameter $\beta_{(s)}$, display virtually complete naïveté, supporting the reduced form results. Estimates of $\beta_{(s)}$ are around 1.03-1.05, statistically indistinguishable from the complete-naïveté case of $\beta_{(s)} = 1$.

The participants display substantially more awareness of others' present bias than of their own. The parameter $\beta_{(o)}$, which captures the participants' beliefs regarding others' present bias, is estimated to be around 0.87, and strongly statistically significantly different from the null of $\beta_{(o)} = 1$. The estimated value of $\beta_{(o)}$ is significantly lower than that of β_s: the bootstrapped t-statistic on the difference between these two parameters is 2.11 without attrited participants, and 2.14 including all participants.

The beliefs-about-others parameter $\beta_{(o)}$ is somewhat higher than the true present bias parameter β, but the difference is not statistically significant. Consistent with the reduced-form results, these parameter estimates indicate that participants are consistently aware of the fact that others will choose to do fewer rounds of the task when the work is imminent, even if they may underestimate the full extent of the difference.

The findings of significant present bias, naïveté about one's own present bias, and awareness of present bias in others are robust to alternative specifications considered in Appendices B.1 and B.2: different baseline levels in ahead-of-time decisions for self versus predicted others and separate estimation of all model parameters for self and others. In these alternative specifications, I find consistent estimates of β around 0.78-0.84 (always significantly different from 1), $\beta_{(s)}$ between 0.99 and 1.01 (never significantly different from 1), and $\beta_{(o)}$ between 0.92 and 0.93 (always significantly different from 1).

Overall, my structural estimates confirm the intuition from the reduced form experimental results: although individuals tend to be naïve about their own present bias, they are more aware of present bias in others.

4.3 Individual Level Estimates

I estimate the structural model individually for each participant, and document two key findings. First, individual-level results confirm the results from pooled analysis: $\beta_{(s)}$ is centered around 1.00, while $\beta_{(o)}$ is centered around 0.93. Second, although the estimates of $\beta_{(s)}$ are perfectly naïve in absolute terms, individual-level estimates of $\beta_{(s)}$ and β are positively correlated, indicating some awareness of relative self-control.

I estimate $\beta^{(i)}$, $\beta_{(s)}^{(i)}$, and $\beta_{(o)}^{(i)}$ for each individual participant i as follows. For participants in Group 3, who face the full set of experimental questions, I estimate the full specification (6) individually for each participant, obtaining individual-level estimates $\beta^{(i)}, \beta_{(s)}^{(i)}$, and $\beta_{(o)}^{(i)}$.

For each participant i in Group 1, who makes predictions about herself but not others, I use her work decisions and predictions regarding her own future work to estimate $\beta^{(i)}$ and $\beta_{(s)}^{(i)}$. Denote the predicted values of individual i's responses by $\hat{e}_{(s)}^{(i)}(t, \tau, w, x, \mathbb{1}_s)$. Then I estimate the following specification:

$$\hat{e}_{(s)}^{(i)}(t, \tau, w, x, \mathbb{1}_s) = \left((\beta^{(i)})^{\mathbb{1}_{t=\tau}}(\beta_{(s)}^{(i)})^{\mathbb{1}_s}\delta^{T-\tau}\phi w\right)^{\frac{1}{\gamma-1}} - x \tag{7}$$

Similarly, for each participant i in Group 2, I use her predictions regarding others' current and future choices, $\hat{e}_{(o)}^{(i)}(t, \tau, w, x)$, to estimate the following specification:

$$\hat{e}_{(o)}^{(i)}(t, \tau, w, x) = \left((\beta_{(o)}^{(i)})^{\mathbb{1}_{t=\tau}}\delta^{T-\tau}\phi w\right)^{\frac{1}{\gamma-1}} - x \tag{8}$$

which yields an individual estimate of the parameter $\beta_{(o)}^{(i)}$.

I then compile individual estimates $\beta^{(i)}$ and $\beta_{(s)}^{(i)}$ from participants making self-predictions (Groups 1 and 3), and individual estimates of $\beta_{(o)}^{(i)}$ from participants making other-predictions (Groups 2 and 3). All individual estimates are constrained to fall between 0.5 and 1.5. The median values and inter-quartile ranges of the three estimated parameters are displayed in Figure 8. These results reflect all participants, including attritors. Without attritors, the median values of the estimated parameters are 0.85 for $\beta^{(i)}$, 1.00 for $\beta^{(i)}_{(s)}$, and 0.94 for $\beta^{(i)}_{(o)}$.

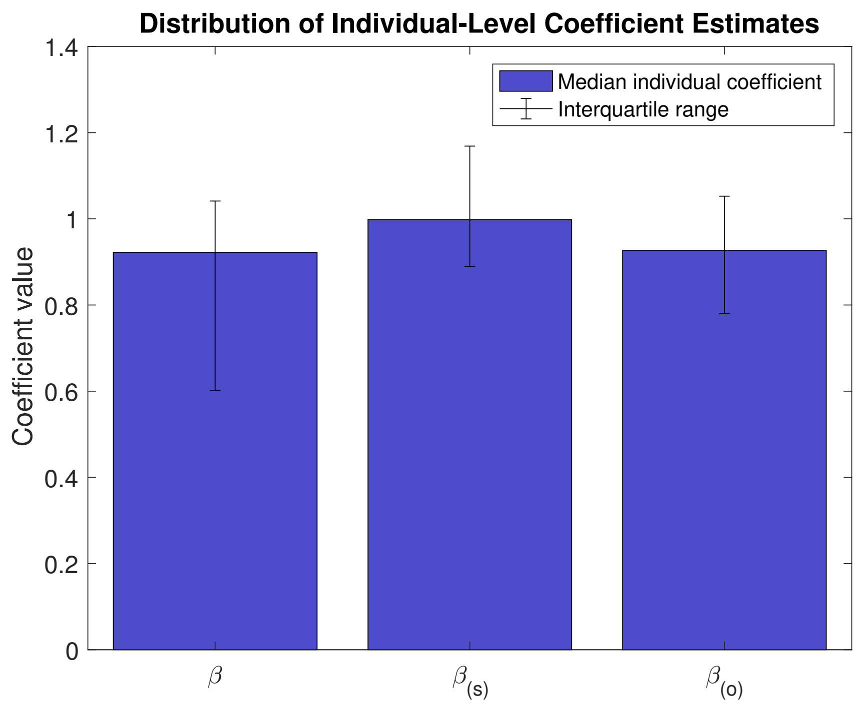


Figure 8: Distribution of the structural parameter estimates across individual experimental participants. The model is estimated separately for each participant i, and the figure presents the median values and interquartile ranges of the estimates of the three key parameters: $\beta^{(i)}$, $\beta^{(i)}_{(s)}$, and $\beta^{(i)}_{(o)}$.

Although individual-level estimates are imprecise due to small amounts of data per participant (e.g., each participant makes only 10 decisions for immediate work), the patterns are consistent with the pooled results. The median value of individual estimates of $\beta^{(i)}$ is 0.92, the median value of estimates of $\beta_{(s)}$ is 1.00, and the median value of individual-level estimates of $\beta_{(o)}$ is 0.93. The mean values of the three estimated parameters are 0.88, 1.02, and 0.94, respectively, consistent with the results in Table 6. The interquartile ranges, however, are quite wide. The 75th percentile of all three estimated parameters is above 1, and the 25th percentile is as low as 0.60 for estimates of $\beta^{(i)}$, 0.89 for $\beta^{(i)}_{(s)}$, and 0.78 for $\beta^{(i)}_{(o)}$.

Interestingly, despite the wedge in median estimates, it is not necessarily the case that predictions regarding randomly matched others are more accurate than self-predictions. I perform the comparison by matching each individual estimate $\beta^{(i)}$ to two predictions: (i) the same individual's self-prediction $\beta^{(i)}_{(s)}$; and (ii) a randomly drawn, with replacement, prediction $\beta^{(j)}_{(o)}$. I then calculate the mean squared error of each set of predictors. The self-predictions yield a mean squared error of 0.1237. I run 1,000 random matches to estimate the mean squared error of the randomly matched other-predictions, and find a median value of 0.1445 and an inter-quartile range of [0.1368, 0.1525]. Thus, it is possible for the self-predictions to actually perform better in matching the present bias parameters $\beta^{(i)}$ than random matching with other-predictions, although this result is sensitive to the exact evaluation criterion used (mean squared error).

The reason for the relatively good performance of self-predictions is the following. While the median value of individual-level $\beta_{(s)}^{(i)}$ is perfectly naïve at 1.00, there is a *positive correlation* between the individual-level parameter estimates. In particular, the correlation between individual-level estimates of $\beta^{(i)}$ and $\beta_{(s)}^{(i)}$ is 0.28. Similarly, a regression of the extent of an individual's self-awareness, $1-\beta_{(s)}^{(i)}$, on the extent of the individual's actual self-control problem, $1-\beta^{(i)}$, yields a coefficient of 0.19, statistically significant at the 1% level. By construction, predictions regarding others are elicited regarding the *average* of others' behavior, and not regarding precise individuals. As a result, although other-predictions are substantially more accurate in general, the average wedge in beliefs may not necessarily compensate for precise knowledge of a given individual.

The last question I pose with the individual-level estimates is whether there is a positive relationship between individuals' beliefs regarding self and others. Conceptually, when forming beliefs regarding unknown others, individuals may project their expectations of their own behavior.[14] This would induce a positive relationship between the estimates of $\beta_{(s)}^{(i)}$ and $\beta_{(o)}^{(i)}$ for those participants who make both sets of predictions. To test this conjecture, I look at individual-level estimates of $\beta_{(s)}^{(i)}$ and $\beta_{(o)}^{(i)}$ from participants in Group 3. Interestingly, I find no positive relationship between beliefs regarding self and others. If anything, correlation between individual-level estimates of $\beta_{(s)}$ and $\beta_{(o)}$ is negative at -0.21. Hence, it does not appear to be the case that my experimental participants project their self-expectations when forming beliefs regarding others.

5 External Evidence: Classroom Experiment

This section presents an intuitive field survey conducted in a classroom, which serves to illustrate the wedge in beliefs regarding one's own and others' present bias in a real-world setting. I first outline the setting and design of the classroom experiment, and then present the results.

[14] A number of studies document that people tend to self-project when forming beliefs regarding others in a variety of domains. For example, Van Boven and Loewenstein (2000) study self-projection of valuations, Van Boven and Loewenstein (2003) consider self-projection of transient drive states, Iriberri and Rey-Bier (2013) find evidence of self-projection of social preferences, and Ludwig and Nafziger (2011) observe self-projection of perceived ability. Madarasz (2012) proposes a model of self-projection of information, and Loewenstein, Moore, and Weber (2006) and Danz, Madarasz, and Wang (2015) offer experimental studies illustrating information projection.

5.1 Design

The classroom experiment is administered to students in an undergraduate financial accounting course (BUS201) at the University of San Francisco. On the first day of class, January 25, 2016, the students are presented with the course syllabus and introduced to the Individual Project that they have to complete for the course, due on May 2, 2016. The project consists of analyzing accounting ratios of a publicly traded company. In order to proceed with the project, students must first choose a company to analyze and confirm that they can download the company's financial statements for the past three years from the Securities and Exchange Commission's website. The students need to email their chosen company and the downloaded financial statements for instructor approval by April 2, 2016. No two students can cover the same company, and approval is granted on a first-come-first-served basis.

Present bias is proxied by the time when the students email the instructor for approval (hereafter referred to as the students' "completion dates"). On the one hand, earlier submission is efficient in that it maximizes the chances of approval (i.e., that no other student has preempted the choice) and leaves more time to work on the project once approved. On the other hand, downloading financial statements carries an immediate effort cost, on which the students might wish to procrastinate. Thus, the more present-biased a given student is, the more likely she is to delay the completion date.

After the project is explained to the students, they are asked to fill out an anonymous, voluntary survey, featuring one or both of the following two questions:

- **Self-Prediction:** As you can see from the syllabus, the deadline for the Individual Project is on May 2, 2016. The last day to submit your chosen company for instructor approval is on April 2, 2016. When do you think you will email your chosen company to the instructor? (Enter a date)

- **Other-Prediction:** As you can see from the syllabus, the deadline for the Individual Project is on May 2, 2016. The last day to submit your chosen company for instructor approval is on April 2, 2016. On average, when do you think your classmates will email their chosen companies to the instructor? (Enter a date)

Each student receives one of four survey versions, distributed randomly among the students:

- Group 1: This version includes the self-prediction question only.

- Group 2: This version includes the other-prediction question only.

- Group 3: This version includes both predictions, with the self-prediction question posed first.

- Group 4: This version includes both predictions, with the other-prediction first.

A total of 57 students attended the class on January 25, 2016, all of whom filled out the voluntary survey. Of these students, 13 were in Group 1, 11 in Group 2, 15 in Group 3, and 18 in Group 4.

5.2 Results

The results of the classroom experiment confirm that the students are significantly more aware of their classmates' procrastination than of their own. While expectations for self are quite overconfident, expectations for others are, on average, correct.

I begin by assessing the differences in the students' predictions about themselves and their classmates graphically. Panel 1 of Figure 9 displays: (i) the distribution of answers among the students predicting for themselves (in dark blue); (ii) the distribution of the students' predictions about their classmates (in light blue); and (iii) the distribution of the actual completion dates (in grey). Only 37% of students making self-predictions expect their completion dates to fall within one week of the deadline (March 27 - April 2, 2016), but a substantially larger proportion (68%) of students expect that others' completion dates would fall, on average, in the last week.

In order to compare the average predictions about self and others, I code predicted completion dates as the number of days before the April 2, 2016 deadline. Thus, for example, a student that predicts that she will email the instructor on March 25, 2016 is coded as making a self-prediction 8 days before the deadline. The average predictions for self and others across survey versions are presented in Panel 1 of Table 7. The average (median) predicted completion date for self is 22.28 days (15.5 days) before the deadline, while the average (median) prediction for others is only 9.07 days (1 day) before the deadline.

To more precisely evaluate the difference between the students' self- and other- predictions, I estimate the following specification:

$$\#DaysBeforeDeadline_i = \alpha + \gamma SelfDummy_i + \epsilon_i, \tag{9}$$

where the response variable $\#DaysBeforeDeadline_i$ denotes the number of days between the prediction i and the deadline (April 2, 2016), and $SelfDummy_i$ is a dummy variable equal to one if prediction i is made about self. In samples including students from Groups 3 and 4, standard errors are clustered by student. The estimates of the difference (coefficient γ) are reported in Panel 2 of Table 7.

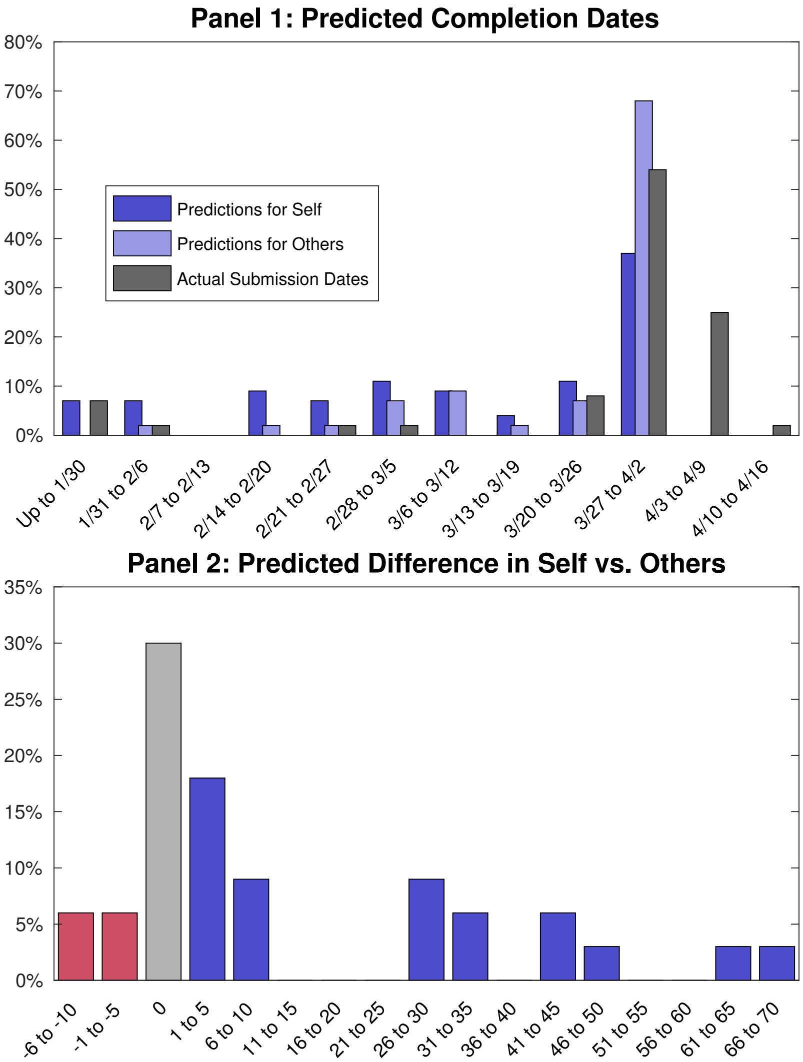


Figure 9: Results from the classroom experiment. **Panel 1** shows the distribution of the predicted assignment completion dates for self and others, as well as the distribution of actual completion dates. **Panel 2** displays the distribution of individual-level differences in predicted completion dates between self and others. For example, a value of 10 corresponds to a student predicting that, on average, others in the class will send the chosen company to the instructor 10 days later than she will.

The average difference between the students' predictions of their own and others' completion dates is 13.21 days. This result is significant at the 1% level when using larger samples (Group 4 or the combined Overall sample), and at the 10% level in smaller samples (Groups

1 & 2 combined, as well as Group 3).

Table 7: Comparison of students' predictions of their own and their classmates' completion dates.
Panel 1: Means, standard errors, and medians of predicted completion dates for self and others. Predicted completion dates are coded as the number of days before the deadline, April 2, 2016.
Panel 2: Differences in predicted completion dates for self and others estimated using the following specification:
$\#DaysBeforeDeadline_i = \alpha + \gamma SelfDummy_i + \epsilon_i$,
where $\#DaysBeforeDeadline_i$ denotes the number of days between the date in prediction i and the deadline, and $SelfDummy_i$ is a dummy variable equal to one if prediction i is made about self. The reported parameter of interest is γ. In samples including Groups 3 and 4, standard errors are clustered by student.
Panel 3: Differences of self- and other-predictions from actual average completion dates. The test compares all predictions of a given type (for self in the left column or for others in the right column) against all actual completion dates.

Panel 1: Summary Statistics						
		Overall	Group 1	Group 2	Group 3	Group 4
Self-prediction	Mean	22.28	22.31	–	21.00	23.33
	SE	*(3.27)*	*(6.58)*	–	*(5.90)*	*(6.01)*
	Median	15.5	18	–	8	15.5
	# Obs	46	13	–	15	18
Other-prediction	Mean	9.07	–	7.91	9.13	9.72
	SE	*(2.15)*	–	*(3.39)*	*(3.57)*	*(5.06)*
	Median	1	–	1	3	1
	# Obs	44	–	11	15	18

Panel 2: Differences in Predictions for Self vs. Others				
	Overall	Group 1 & Group 2	Group 3	Group 4
Difference	**13.21**	**14.40**	**11.87**	**13.61**
SE	*(3.24)*	*(7.42)*	*(6.15)*	*(4.29)*

Panel 3: Diff in Predictions from Actual Completion Dates		
	Self-prediction	Other-prediction
Difference	**14.82**	**1.61**
SE	*(3.94)*	*(3.36)*

Just as in the online laboratory experiment, the predictions are independent of either the set of questions asked or the order in which they are asked. In particular, posing the two questions side by side to the same students (responses from Groups 3 and 4) yields the same results as asking different groups of students to make the two sets of predictions (responses from Groups 1 an 2). Similarly, varying the order in which the students in Groups 3 and 4 see the two questions does not materially affect the results.

For the students who make both sets of predictions, I observe the distribution of the individual-level differences in predicted dates. This distribution, presented in Panel 2 of Figure 1, displays the incidence of individual students being more optimistic about themselves, being more optimistic about others, or holding identical beliefs about themselves and others. While a large portion of students (30%) make the exact same prediction for themselves as for their average classmates, the clear majority (58%) expect others to email the instructor later than they will. Only 12% of the respondents expect themselves to email the instructor later than their classmates. Thus, the individual-level results confirm the patterns of the pooled analysis: the students expect themselves to display, on average, less present bias (i.e., have earlier completion dates) than their peers.

Which set of predictions is more correct? A cursory examination of the distributions in Panel 1 of Figure 9 indicates that the more critical expectations about others (in light blue) more closely match the real distribution of completion date (in grey). In fact, a sizable proportion of the students (26%) email the instructor for approval a few days *after* the April 2, 2016 deadline.

The notion that beliefs about others more closely match real completion dates is confirmed by a statistical comparison of the average predictions from students making self- and other- predictions against the average actual completion dates, presented in Panel 3 of Table 7. The students' predictions of their own completion dates are, on average, a full two weeks off from actual completion dates. The difference between self-predictions and actual completion dates is statistically significant at the 1% level. Predictions about others, however, are remarkably spot-on. The average difference between the students' predictions of when their classmates would email the instructor and actual completion dates is economically negligible (1.61 days) and statistically indistinguishable from zero.

Altogether, the results of the classroom experiment illustrate that the wedge in beliefs documented in my laboratory experiment is operative in a real-world setting: the classroom. Among college students, there is a strong tendency to anticipate last-minute assignment completion from their classmates, but not from themselves. A more detailed discussion of applications of the documented wedge in beliefs is presented in the next section.

6 Applications

In this section, I offer examples of settings featuring interactions among present-biased individuals, where beliefs regarding others' present bias could affect equilibrium outcomes. I split these illustrative scenarios into three broad categories: competitive environments, collaborative group decision making, and principal-agent problems.

6.1 Competitive Environments

Beliefs regarding peers are instrumental for behavior in competitive settings. Under competitive incentive structures such as tournaments,[15] the optimal level of expended effort depends on one's expectations regarding the behavior of others.

Ex ante participation. Relatively more optimistic beliefs regarding one's own versus others' present bias makes entering competitive incentive schemes appear more attractive. For example, consider an employee faced with a tournament incentive contract. The wedge in beliefs makes the employee overestimate his future effort relative to his peers, inflating perceived chances of receiving the prize. This relaxes his participation constraint, potentially resulting in acceptance of contracts with negative net present value, and can lead to exploitative situations. O'Donoghue and Rabin (1999b) showcase the potential for firms to exploit naïve workers with individual performance-based pay. The wedge in beliefs regarding self versus others creates scope for further exploitation through relative performance contracts.

Ex post efficacy. However, asymmetric naïveté can also reduce the ex-post efficacy of tournaments as incentive devices. For example, consider an employee receiving compensation based on his relative performance. Due to asymmetric naïveté, he expects his peers to procrastinate, but holds overconfident beliefs regarding his future self. As a result, the employee can underestimate the cost of delaying his work, and expend less effort initially. Similar logic applies in contexts such as queueing, trading in financial markets in response to new information, or otherwise expending effort to secure a limited resource in a competitive marketplace. Overall, the wedge in beliefs regarding own vs. others' present bias exacerbates procrastination in competitive environments.

6.2 Group Decision Making

Beliefs regarding others play a role not only in competitive situations, but also in collaborative settings. Below, I outline a few examples of joint decision making where each party's expectations regarding the other's present bias can influence equilibrium behavior.

Household consumption. A number of studies document that individuals display present bias in their day-to-day household decisions including credit card usage (Meier and Sprenger, 2010), saving for retirement (see Laibson, Repetto, and Tobacman, 1998; Choi, Laibson, and Madrian, 2011), and exercise (DellaVigna and Malmendier, 2006). Asymmetrically naïve individuals are likely to make more sophisticated decisions on behalf of their

[15] See Lazear and Rosen (1981) and Green and Stokey (1983), among others.

partners or children than on their own behalf. This predicts higher willingness to enroll in commitment devices such as savings accounts when the decision is made jointly with a spouse versus by oneself, as well as potential gains from joint time management. The wedge in beliefs can also increase efficiency gains from group commitment devices such as weight loss programs and group-lending in microcredit markets.

Teamwork. Understanding of others' present bias can also serve as a valuable disciplining and commitment device for teams in the workplace (Gans and Landry, 2016). Fedyk (2015) highlights one channel through which team assignments can improve performance in the face of asymmetric naïveté. When a naïve present biased individual is assigned a task to be completed within a predetermined amount of time, her overoptimistic beliefs regarding her future present bias cause her to overestimate the option value of postponing the task until later. By contrast, awareness of each other's present bias makes each member of a team more willing to commit to completing the work early on, so as not to leave scope for her teammate's future procrastination. More generally, awareness of other's present bias makes commitment devices such as preplanned meetings, milestones, and deadlines more attractive to teams than to standalone workers.

Collaboration and delegation. However, asymmetric naïveté can also have detrimental effects on performance. In arenas where teamwork is not mandated, the wedge in beliefs regarding own versus others' present bias makes teamwork appear (erroneously) less appealing. This can lead to inefficiently low levels of collaboration among peers. Similarly, in existing teams, asymmetric naïveté can create the temptation for an overconfident individual to take on too much, displaying suboptimally low levels of delegation of tasks across teammates.

6.3 Principal-Agent Settings

Most principal-agent models featuring present bias assume a rational and omniscient principal interacting with present-biased agent(s).[16] However, the extent to which firms, employers, and governments can correctly assess consumers', employees', and individuals' present bias is an empirical question. The wedge in beliefs documented in the present paper gives some credence to the standard theoretical assumption of omniscient principals. Below I summarize the implications in the classroom and in the workplace.

[16]See, for example, DellaVigna and Malmendier (2004) and O'Donoghue and Rabin (2006). Similarly, Heidhues and Kőszegi (2010) assume that firms can directly observe either consumers' β or $\hat{\beta}$ parameters, as well as the structural relationship between the two. O'Donoghue and Rabin (1999b) assume that the principal knows the agent's present-bias parameter β and model the principal's uncertainty regarding the agent's idiosyncratic completion costs.

Classroom. A common feature of classroom instruction is deadlines, including homework assignments, in-class presentations, and intermediate exams. Without these mechanisms, a student's naïveté regarding his own present bias would prompt him to set overly flexible deadlines (Ariely and Wertenbroch, 2002). Yet a teacher's ability to assign work in a way that maximizes her students' effort and performance hinges on her understanding of the students' present bias. The present paper indicates that naïveté regarding present bias is asymmetric. In the teacher-student setting, this implies that teachers hold more critical beliefs regarding their students' present bias, and hence impose more effective deadlines.

Workplace management. Analogous to the commitment devices offered by teachers in the classroom, managers serve a similar function in organizations. In a field experiment conducted at a Colombian bank, Cadena, Cristea, Delgado-Medrano, and Schoar (2011) show that greater paternalistic incentives, such as goal reminders and managerial monitoring, lead to not only superior on-the-job performance and earnings, but also higher ex post employee satisfaction and lower stress levels. As Laibson (2018) highlights, individuals do not explicitly seek out more paternalistic and restrictive workplaces (consistent with naïveté regarding own present bias), but work environments invariably provide restrictions such as intermediate deadlines and monitored attendance (consistent with general awareness of others' present bias). The wedge in beliefs creates the scope for such private paternalism: through their more critical awareness of others' present bias, managers serve as more effective organizational devices than employees' own planning.

Altogether, beliefs regarding others' present bias can inform our understanding of equilibrium outcomes across competitive, collaborative, and hierarchical environments. The examples above discuss several potential applications, but the list is by no means exhaustive, since present bias has been shown to be operative in a variety of domains, and strategic interactions feature in multiple types of economic decision-making.

7 Conclusion

This paper investigates whether individuals are aware of present bias in others. Both the online laboratory experiment and the field survey in the classroom reveal a wedge in beliefs: individuals are fairly naïve about their own present bias, but anticipate present bias in others. This finding is robust to incentivizing the predictions with monetary payments, and to asking the two sets of predictions – about self and about others – to the same experimental participants versus separately to different groups of participants. The wedge in beliefs explored in this paper opens two avenues for future work.

First, further investigation of the documented wedge in beliefs can shed light on the mechanisms of belief formation. My findings indicate that present bias is subject to relative overconfidence akin to that documented in several other domains.[17] In addition, my results support the notion of bias blind spots documented in the social psychology literature: that individuals are, in general, more perceptive of others' biases than of their own.[18] Probing further into how beliefs regarding one's own and others' present bias evolve depending on the setting, task, or experience can help shed light on the extent to which the documented wedge in beliefs reflects motivated thinking, blindspots, or other frictions.

Second, the documented wedge in beliefs lays the foundations for understanding interactions between present-biased individuals in the workplace, in the classroom, in households, and in markets. Differential awareness of one's own and others' present bias is likely to impact how groups of present-biased individuals schedule their joint work, seek external commitment devices, or evaluate their own and their peers' performance. Investigating these effects, both theoretically and empirically, could constitute a fruitful avenue for future research.

References

[1] Acland, D. and M. R. Levy (2015). "Naiveté, projection bias, and habit formation in gym attendance." *Management Science,* 61.1: 146-160.

[2] Alicke, M. D. (1985). "Global self-evaluation as determined by the desirability and controllability of trait adjectives." *Journal of Personality and Social Psychology,* 49.6: 1621-1630.

[3] Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutstrom (2008). "Eliciting risk and time preferences." *Econometrica*, 76.3: 583-618.

[4] Ariely, D. and G. Loewenstein (2006). "The heat of the moment: The effect of sexual arousal on sexual decision making." *Journal of Behavioral Decision Making*, 19.2: 87-98.

[5] Ariely, D. and K. Wertenbroch (2002). "Procrastination, deadlines, and performance: Self-control by precommitment." *Psychological Science*, 13.3: 219-224.

[6] Augenblick, N., M. Niederle, and C. Sprenger (2015). "Working over time: Dynamic inconsistency in real effort tasks." *Quarterly Journal of Economics,* 130.3: 1067-1115.

[7] Augenblick N. and M. Rabin (2018). "An experiment on time preference and misprediction in unpleasant tasks." *Review of Economic Studies,* forthcoming.

[17] For example, Svenson (1981) documents overconfidence regarding driving skills, while Weinstein (1980) finds overconfidence and overoptimism about a host of potential life events. Alicke (1985) documents that participants deem positive (negative) adjectives to be more (less) characteristic of themselves than of their average peer.

[18] See, for example: Pronin, Lin, and Ross (2002), Ehrlinger, Gilovich, and Ross (2005), and West, Meserve, and Stanovich (2012).

[8] Bisin, A., and K. Hyndman (2014). "Present-bias, procrastination and deadlines in a field experiment." NBER Working Paper No. 19874.

[9] Cadena, X., A. Schoar, A. Cristea, and H. Delgado-Medrano (2011). "Fighting procrastination in the workplace: An experiment." NBER Working Paper No. 16944.

[10] Choi, J., D. Laibson, and B. Madrian (2011). "$100 bills on the sidewalk: Suboptimal investment in 401(k) plans." *Review of Economics and Statistics,* 93(3): 748-763.

[11] Conlin, M., T. O'Donoghue, and T. J. Vogelsang (2007). "Projection bias in catalog orders." *American Economic Review,* 97.4: 1217-1249.

[12] Danz, D., K. Madarász, and S. W. Wang (2015). "Anticipating information projection: An experimental investigation." Working paper.

[13] DellaVigna, S. and U. Malmendier (2004). "Contract design and self-control: Theory and evidence." *Quarterly Journal of Economics*, 119.2: 353-402

[14] DellaVigna, S. and U. Malmendier (2006). "Paying not to go to the gym." *American Economic Review*, 96.3: 694-719.

[15] DellaVigna, S. and M. D. Paserman (2005). "Job search and impatience." *Journal of Labor Economics*, 23.3: 527-588.

[16] Ehrlinger, J., T. Gilovich, and L. Ross (2005). "Peering into the bias blind spot: People's assessments of bias in themselves and others." *Personality and Social Psychology Bulletin*, 31.5: 680-692.

[17] Fahn, M. and H. Hakenes (2014). "Teamwork as a self-disciplining device." CESifo Working Paper No. 5131.

[18] Fedyk, A. (2015). "Overcoming overconfidence: Teamwork and self-control." Working paper, Harvard University.

[19] Gans, J. S. and P. Landry (2016). "Procrastination in teams." NBER Working Paper No. w21891.

[20] Gottlieb, D. (2008). "Competition over time-inconsistent consumers." *Journal of Public Economic Theory*, 10.4: 673-684.

[21] Green, J. and N. Stokey (1983). "A comparison of tournaments and contracts." *Journal of Political Economy*, 91.3: 349-364.

[22] Gul, F. and W. Pesendorfer (2001). "Temptation and self-control." *Econometrica*, 69.6: 1403-1435.

[23] Heidhues, P. and B. Kőszegi (2010). "Exploiting naïveté about self-control in the credit market." *American Economic Review*, 100.5: 2279-2303.

[24] Herweg, F. and D. Müller (2011). "Performance of procrastinators: On the value of deadlines." *Theory and Decision*, 70.3: 329-366.

[25] Iriberri, N. and P. Rey-Biel (2013). "Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do?" *Quantitative Economics,* 4.3: 515-547.

[26] Kaufmann, M. (2017). "Projection bias in effort choices." Working paper, Central European University.

[27] Laibson, D. (1997). "Hyperbolic discounting and golden eggs." *Quarterly Journal of Economics*, 112.2: 443-477.

[28] Laibson, D. (2018). "Private paternalism, the commitment puzzle, and model-free equilibrium." *American Economic Review: Papers & Proceedings.*

[29] Laibson, D., A. Repetto, and J. Tobacman (1998). "Self-control and saving for retirement." *Brookings Papers on Economic Activity*, 1: 91-196.

[30] Laibson, D., A. Repetto, and J. Tobacman (2008). "Estimating discount functions with consumption choices over the lifecycle." NBER Working Paper No. 13314.

[31] Lazear, E. and S. Rosen (1981). "Rank order tournaments as optimum labor contracts." *Journal of Political Economy*, 89.5: 841-864.

[32] Loewenstein, G. and D. Adler (1995). "A bias in the prediction of tastes." *Economic Journal*, 105.431: 929-937.

[33] Loewenstein, G., D. Moore, and R. Weber (2006). "Misperceiving the value of information in predicting the performance of others." *Experimental Economics*, 9.3: 281-295.

[34] Loewenstein, G., T. O'Donoghue, and M. Rabin (2003). "Projection bias in predicting future utility." *Quarterly Journal of Economics*, 118.4: 1209-1248.

[35] Ludwig, S. and J. Nafziger (2011). "Beliefs about overconfidence." *Theory and Decision*, 70.4: 475-500.

[36] Madarasz, K. (2012). "Information projection: Model and applications." *Review of Economic Studies*, 79.3: 961-985.

[37] McClure, S.M., D. Laibson, G. Loewenstein, and J. D. Cohen (2004), "Separate neural systems value immediate and delayed monetary rewards." *Science*, 306.5695: 503-507.

[38] Meier, S. and C. Sprenger (2010). "Present-biased preferences and credit card borrowing." *American Economic Journal: Applied Economics,* 2.1: 193-210

[39] O'Donoghue, T. and M. Rabin (1999a), "Doing it now or later." *American Economic Review*, 89.1: 103-124.

[40] O'Donoghue, T. and M. Rabin (1999b). "Incentives for procrastinators." *Quarterly Journal of Economics*, 114.3: 769-816.

[41] O'Donoghue, T., and M. Rabin (2006). "Optimal sin taxes." *Journal of Public Economics,* 90.10-11: 1825-1849.

[42] Pronin, E., D. Y. Lin, and L. Ross (2002). "The bias blind spot: Perceptions of bias in self versus others." *Personality and Social Psychology Bulletin,* 28.3: 369-381.

[43] Read, D. and B. Van Leeuwen (1998). "Predicting hunger: The effects of appetite and delay on choice." *Organizational Behavior and Human Decision Processes*, 76.2: 189-205.

[44] Shui, H. and L. Ausubel (2005). "Time inconsistency in the credit card market." Working Paper, University of Maryland.

[45] Skiba, P. and J. Tobacman (2009). "Payday loans, uncertainty and discounting: Explaining patterns of borrowing, repayment, and default." NBER Working Paper No. 14659.

[46] Solnick, J.V., C. H. Kannenberg, D. A. Eckerman, and M. B. Waller (1980). "An experimental analysis of impulsivity and impulse control in humans." *Learning and Motivation*, 11.1: 61-77.

[47] Svenson, O. (1981). "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica,* 47.2: 143-148.

[48] Tanaka, T., C. F. Camerer, and Q. Nguyen (2010). "Risk and time preferences: Linking experimental and household survey data from Vietnam." *American Economic Review*, 100.1: 557-571.

[49] Tversky, A. and D. Kahneman (1974). "Judgment under uncertainty: Heuristics and biases." *Science,* 185.4157: 1124-1131.

[50] Van Boven, L. and G. Loewenstein (2000). "Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect." *Journal of Personality and Social Psychology*, 79.1: 66-76.

[51] Van Boven, L. and G. Loewenstein (2003). "Social projection of transient drive states." *Personality and Social Psychology Bulletin*, 29.9: 1159-1168.

[52] Weinstein, N. D. (1980). "Unrealistic optimism about future life events." *Journal of Personality and Social Psychology*, 39.5: 806-820.

[53] West, R. F., R. J. Meserve, and K. E. Stanovich (2012). "Cognitive sophistication does not attenuate the bias blind spot." *Journal of Personality and Social Psychology,* 103.3: 506-519.