Games and Economic Behavior 119 (2020) 339–357

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Present-bias, procrastination and deadlines in a field experiment ☆

Alberto Bisin [a,b,c,1], Kyle Hyndman [d,e,*,2]

[a] New York University, United States of America
[b] CESS, United States of America
[c] NBER, United States of America
[d] University of Texas at Dallas, United States of America
[e] CESS, United States of America

ARTICLE INFO

Article history:
Received 18 August 2018
Available online 10 December 2019

Keywords:
Procrastination
Deadlines
Commitment
Experiment

ABSTRACT

We study procrastination in the context of a field experiment involving students who must exert costly effort to complete certain tasks by a fixed deadline. We document a robust demand for commitment, in the form of self-imposed deadlines. On the other hand, deadlines do not increase completion rates in our experiment. Furthermore, while we find that present-bias is widespread in the sample, and present-biased students procrastinate in single task treatments, we find that they successfully manage to self-control in repeated task treatments. Finally, we find evidence that students do not set deadlines optimally and that deadlines may hurt them, due to various behavioral components of students' anticipation formation mechanisms; specifically, partial naïveté at the deadline setting stage and over-confidence about the ability to complete the task and to persevere on a task after a failed attempt.

1. Introduction

Procrastination is generally defined in the psychological literature as the practice of putting off impending tasks to a later time even when such practice results in "counterproductive and needless delay" (see e.g., Schraw et al., 2007). The qualification that delay be counterproductive and needless is important. Delay may represent an optimal strategy in an environment in which the cost of effort evolves over time, when waiting for the best moment to complete a task. Procrastination is then typically construed in psychology and economics as the result of a present-bias in preferences, on

☆ We thank Anne Stubing, Aditya Bhandari, Margaret Ford, Taylor McBride, and Daniel Vaughan for their excellent help with the experiments; Nicholas DePinto, Chardonnay Phelan, and Severine Toussaert for their research assistance; and Anwar Ruff for his programming expertise. We thank the Center for Experimental Social Science for their generous use of their lab. We have received very helpful comments from Gary Charness, Chetan Dave, John Duffy, Guillaume Frechette, Anett John, David Laibson, David Levine, Johanna Mollerstrom, Andy Schotter, and many seminar participants. Thanks also to Dan Ariely for kindly providing us with the data regarding his seminal experiment on procrastination with Klaus Wertenbroch. Finally, we are thankful to the Co-Editor (Gary Charness) and three anonymous referees for their very helpful comments. Financial support from Southern Methodist University (Hyndman) and from the Center for Experimental Social Science (Bisin & Hyndman) and the C.V. Starr Center for Applied Economics (Bisin) is gratefully acknowledged.

* Corresponding author.
E-mail addresses: alberto.bisin@nyu.edu (A. Bisin), KyleB.Hyndman@utdallas.edu (K. Hyndman).
[1] Department of Economics, New York University, 19 W 4th Street, 6th Floor, New York, NY 10012.
[2] Naveen Jindal School of Management, University of Texas at Dallas, 800 W. Campbell Rd. (SM31), Richardson, TX 75080.

https://doi.org/10.1016/j.geb.2019.11.010
0899-8256/

account of which agents delay doing unpleasant tasks that they themselves wish they would do sooner (O'Donoghue and Rabin, 1999a).

In this paper we experimentally study procrastination in students' academic work – a context procrastination appears widespread in. Solomon and Rothblum (1984) find that at least 46% of college students consider themselves serious procrastinators; Steel (2007) finds that between 80% and 95% of college students regularly procrastinate when performing academic tasks. Novarese and Giovinazzo (2013) also studies university administration data concluding that lack of student promptness in enrollment is negatively correlated with academic achievement, a finding which could be interpreted as due to procrastination. Indeed several recent field experiments on procrastination, e.g., Ariely and Wertenbroch (2002), Burger et al. (2011), have focused on students' homework activity. We design and conduct a framed field experiment in which students must exert costly effort to perform a certain number of tasks by a fixed deadline for a monetary payment after completion of each task. Subjects choose when to complete the task(s), if ever, over the course of their normal daily activities, and balance the requirement of the experimental task(s) with the various demands on their time (i.e., academic, employment, leisure).[3] We conceptualize these demands on time as an effort cost associated with each task.

In a dynamic choice context like the one we study, students with a present-bias might adopt various internal (psychological) and/or external self-control mechanisms to avoid procrastinating on the task(s). Internal mechanisms include mental deadlines, cues, and anticipatory planning. External mechanisms include binding self-imposed deadlines and voluntary exposure to social pressure. We explicitly study the role of binding deadlines in affecting procrastination. Furthermore, we infer on the role of internal mechanisms comparing students' behavior when faced with a single task versus multiple repeated tasks, the latter of which have been shown to induce self-regulatory behavior (see e.g., Baumeister et al., 1994; Kuhl and Beckmann, 1985 and Gollwitzer and Bargh, 1996). In the experiment we conduct, subjects must complete either one or up to three tasks over a period of one or two weeks. For each task that is completed by the relevant deadline, a subject receives a fixed monetary reward. In all treatments, subjects face a final, externally imposed deadline; however, different treatments vary regarding whether or not there are binding *intermediate* deadlines as well as whether subjects can commit themselves to an earlier deadline.

Our experiment provides us with several interesting findings. In particular, we document that a significant subset of the students, when given the opportunity, do self-impose a binding deadline. This suggests a *robust demand for commitment* on their part, which is typically interpreted as evidence of present-bias. On the other hand, treatments with either exogenous or endogenous deadlines are not associated with more frequent or earlier completion. In other words, *deadlines do not* appear to *increase task completion rates*. In order to more fully understand the role of deadlines in students' task completion behavior, we then identify which students are present-biased. To this end, we exploit our initial survey data to look at the psychological correlates of behavior and deadline setting. We find that, notably, it is subjects who report themselves to be *less* conscientious who are more likely to set a binding deadline and who complete tasks later. This suggests that indeed *present-bias*, proxied by lack of conscientiousness, *is relatively widespread* in our sample and that it represents an important driver of the demand for commitment. Having identified present-biased students, we are able to compare their distribution of task completions with that of non-present-biased students. In this respect results are particularly striking and novel: while *present-biased students procrastinate in 1T treatments*, *in the repeated task treatments (3T) they successfully manage to self-control*; that is, exponential and present-biased students behave very similarly. We interpret this result as evidence of the ability of students to successfully adopt internal self-control mechanisms when facing repeated tasks.

Our data on deadline-setting suggests that, interpreting the large demand for deadlines we observe in the experiment simply as due to present-bias might not be completely correct. Most importantly, this explanation is at odds with the fact that present-biased students self-impose binding deadlines even though they successfully exercise self-control in No Deadlines treatment with repeated tasks. Indeed, in the case of repeated-task treatment, if students *correctly* anticipated their ability to self-control, then they should not have self-imposed deadlines. This is because deadlines are just as costly for a decision maker who can perfectly self-control as they are for a non-present-biased decision maker. Our data suggest two kinds of complementary behavioral anticipation formation mechanisms. First, students under-estimate their ability to exercise self-control (at least in repeated tasks), a form of *partial naïveté at the deadline setting stage*. Second, students are over-confident about their ability to complete the task(s); in particular they over-estimate their ability to persevere on the current and future tasks after a failed attempt. We also present indirect evidence that over-confidence and lack of perseverance may be distinct psychological characteristics from present-bias.

These behavioral mechanisms are complementary in the sense that both might make students believe that deadlines are more beneficial than they actually are, leading them to self-impose deadlines, which end up harming them. The failure of deadlines to boost completion rates appears related to other interesting behavioral regularities in the data. For instance, we see that many subjects, when faced with a deadline, begin working on the task relatively close to the deadline and, because mistakes frequently occur, they often do not successfully complete tasks that they attempted. Moreover, once a subject gives up on one task, she is very likely to effectively quit the experiment and make no attempt of later tasks.

Fundamentally, deadlines have three distinct effects: (i) they provide commitment; (ii) they impose a cost in terms of the option value of waiting for a lower cost of task completion; (iii) they amplify the cost associated with over-confidence and lack of perseverance. We argue that our data suggest that partial naïveté and over-confidence might have a negative effect

[3] In this paper, we use the words 'students' and 'subjects' interchangeably.

on present-biased students who are able to self-control (and possibly exponential students who might not fully realize they are not present-biased). This would be especially the case in repeated-task experiments, where self-control is widespread in our data.

Our paper contributes to the rich experimental literature in psychology and economics which has first motivated and then supported the behavioral economics of intertemporal choice. This literature has provided strong evidence for present-bias (see, e.g., Ainslie, 1992, 2001; Loewenstein and Prelec, 1992; Frederick et al., 2002). Also, ample evidence of subjects adopting external commitment devices such as binding self-imposed deadlines is obtained both in the lab and in the field. With regards to lab experiments, Trope and Fishbach (2000), Casari (2009), Houser et al. (2010), Toussaert (2018) experimentally study different commitment mechanisms in different contexts, invariably documenting a large demand for commitment. As for field evidence, similar results are obtained by Schilbach (2019), Giné et al. (2010).[4]

Most related to our paper are two recent papers on procrastination in the context of students' academic work, Ariely and Wertenbroch (2002), Burger et al. (2011). Results are somewhat mixed, in that Ariely and Wertenbroch (2002) finds positive effects of deadlines as external commitment devices, while Burger et al. (2011) do not. We follow these papers in the general approach of examining the role of deadlines and, like Ariely and Wertenbroch (2002), in exploiting the demand for commitment to identify present-bias and possibly procrastination. However, we diverge from them in several important elements of the experimental design, as well as in the data we have available to explore. We shall discuss in some detail these differences in Section 4.5, where we try and rationalize these results and ours.

2. Experimental design

We conduct two distinct sets of experiments. In the first, students have one week to complete a single task. We distinguish two treatments corresponding to two different intermediate (before the natural end-of-experiment) deadline scenarios: No Deadline and Endogenous (i.e., self-imposed) Deadlines. We call these the 1T(ask) treatments. In the second set of experiments, students have two weeks to complete three tasks, with three different treatments corresponding to different intermediate deadline scenarios: No Deadlines, Exogenous Deadlines and Endogenous Deadlines. We call these the 3T(ask) treatments.

In the 1T treatments, subjects are paid $20 if they successfully complete the task, while in the 3T treatments, subjects are paid $15 for each task successfully completed by the relevant deadline. In what follows we describe the experimental procedures we use for the 3T treatments. Identical procedures are used for the 1T treatments.

2.1. Phase 1: the lab-based component

Each session begins with a lab-based component in which students read the instructions for their treatment and are given a user name and password in order to gain access to the web-based experimental software. The instructions outline the nature of the tasks and provided a screen shot showing subjects a sample task (as in, e.g., Fig. 1). The instructions also tell students the nature of any deadlines that they face.

After reading the instructions (a sample of which can be found in Appendix A.1), students log on to the experimental software and are reminded of their deadlines for each task. For students in the No Deadlines treatment, all tasks have a deadline set at the end of the experiment; i.e., two weeks after coming into the lab for 3T and one week for 1T. For students in the Exogenous Deadlines treatment, each of the three tasks has a different deadline; deadlines are evenly spaced, with the deadline for task 3 being at the end of the experiment. Students in the Endogenous Deadlines treatment are able to choose an intermediate deadline for each of the three tasks. The latest deadline that students could set is the end of the experiment.

After observing or choosing their deadlines, in the lab, students answer a series of survey questions. The survey asks about their (work, academic and social) schedules for the two-week duration of the experiment. It also asks students to report their subjective expectation (in probability form) of completing 0, 1, 2 or all 3 tasks. Finally, the survey asks a number of questions designed to gauge students' perceptions about several of their own psychological characteristics, like reliability, punctuality, organization, *etc*. Appendix A.2 contains the exact survey questions subjects were asked in both the 1T and 3T treatments.[5]

This component of the experiment is conducted at the Center for Experimental Social Science (C.E.S.S.) at New York University and lasted between 30 and 45 minutes. At the end of this phase, students are given a $10 participation fee.

We should remark that although subjects were provided with a sample task, they were not given the opportunity to complete a practice task in the lab. The instructions did, however, provide some guidance on how subjects should complete the task. Specifically, the instructions stated:

[4] Other field evidence examines voluntary exposure to social pressure (see Mahajan and Tarozzi, 2011; Schwartz et al., 2014, 2012; Bryan et al., 2010, for examples and discussion). A few studies have also shown (cf., Thaler and Benartzi, 2004; Ashraf et al., 2006 and Duflo et al., 2011) that products with certain commitment features lead to higher savings; that workers are willing to choose dominated contracts as a commitment device to increase their productivity (Kaur et al. (2010)). See Bryan et al. (2010) for a comprehensive survey of the literature on commitment and self-control.

[5] The only difference in survey questions between the 1T and 3T treatments were with respect to beliefs. For the 1T treatment, the only relevant belief is the probability of completing the single task, while for the 3T treatments, we also asked subject to predict their likelihood of finishing 0, 1, 2 or 3 tasks.

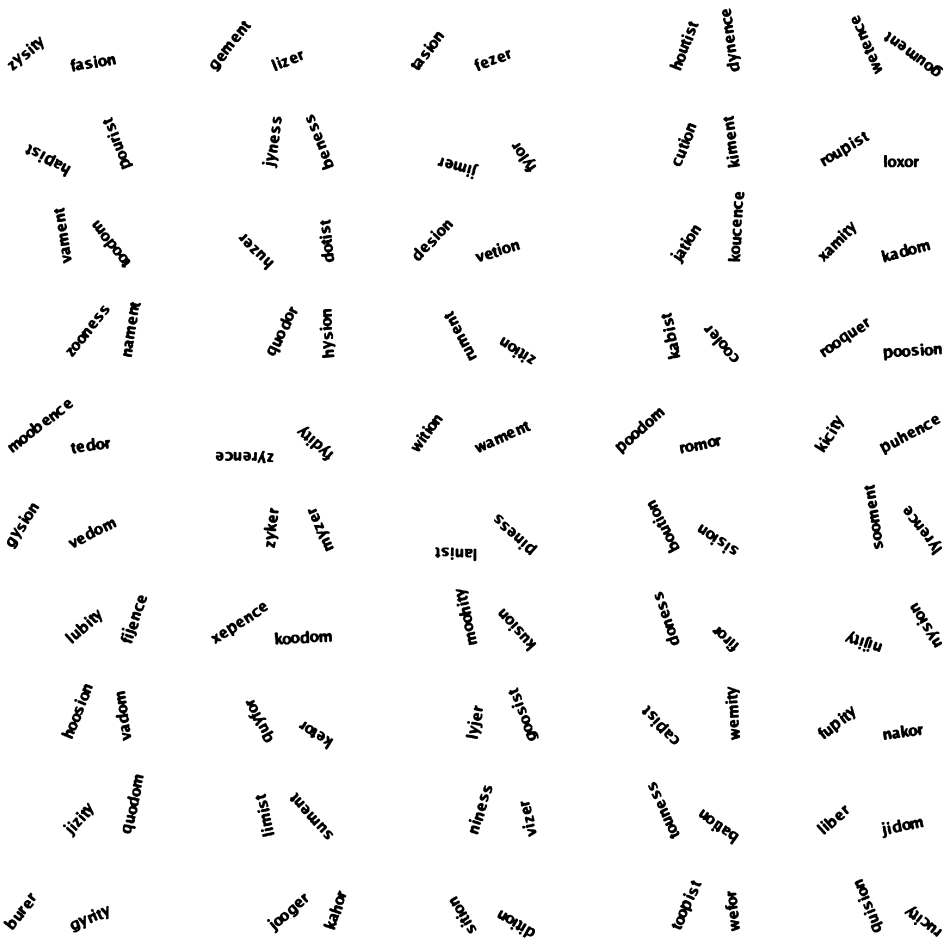


Fig. 1. A sample task.

Of course, there are many different ways that one might wish to approach this task. One way that we have found to work reasonably well is to print the screen containing the words, enter the words into a spreadsheet application such as Excel (Microsoft), Numbers (Apple) or Calc (Open Office), use the sort command and then enter the words in the appropriate order through the experimental interface. If you are careful with this method, then it should be possible to complete each task in 1 hour or less.

In practice, most subjects required less than an hour to complete tasks.[6] Subjects who took longer to complete the task did so, likely, because they made mistakes, which are difficult to identify and correct.

2.2. Phase 2: the experiment

Upon completing the first component of the experiment, students leave the C.E.S.S. lab and are free to work on the tasks at any time they wish. To do so, students log on to a website using their user name and password. Upon logging in, they are issued a list of words for the current task and are asked to list them in alphabetical order. In order to simulate as best as possible a stopping time problem, once a list of words is given, students have to alphabetize the list within the lesser of 2 hours and the time until the task deadline. Failing to do so implies that a new list of words is issued if time remains; if no time remains before that task's deadline, students are automatically taken to the next task. Additionally, each time students refresh the browser or log into the software, a new list of words is issued.

If a student submits an incorrectly alphabetized list, the software sends a message alerting her of the existence of at least one mistake in the submitted list, without any indication about the number of mistakes or their position within the

[6] It is difficult to give a precise estimate of how long it took subjects to complete the task. While some subjects were able to complete a task in as few as 20 minutes, other subjects took nearly the full two hours. Moreover, some subjects attempted a task, gave up and came back later to try again to complete the task.

Table 1
Summary of the various treatments and sessions.

Treatment	Sessions	Intermediate deadlines	Tasks	Number of subjects
1T-None	1	None	1	46
1T-Endog.	1	Endogenous	1	35
3T-None	3	None	3	23, 24, 14
3T-Exog.	3	Exogenous	3	21, 24, 24
3T-Endog.	3	Endogenous	3	21, 24, 22

Note: All 1T treatments were conducted in the middle of the semester and consisted of 150 words. In the 3T treatments, sessions 1 and 3 were conducted in the middle of the semester, while session 2 was conducted at the end of the semester. Sessions 1 and 2 had 150 word tasks, while session 3 had 200 words.

list. If a student submits a correctly alphabetized list, she is immediately taken to the next task, which she can work on if she so chooses.

Each task that is successfully completed by the relevant deadline generates a payment of $15, via petty cash vouchers mailed to students.[7] In particular, all tasks that are completed by 1:00PM on a given day are processed for payment that same day. Tasks completed after 1:00PM or on weekends are processed the next weekday.

2.3. Phase 3: post-experiment survey

Upon completion of the third task, or after the end of the experiment, students are asked to complete a post-experiment survey. The purpose of this is to gain information on any unanticipated shocks that they may have faced during the field component of the experiment.

2.4. Different sessions

In Table 1 we summarize the details of our experimental sessions. In the 3T treatments, we conducted three different sessions at different times. The first two sessions were conducted during the Spring semester of 2010, while the third session took place during the Spring semester of 2011. Importantly, however, for each session, each of the different deadline treatments were conducted at the same time. The second and third sessions were aimed at adding exogenous variation in the cost of completing tasks. In particular, Session 2 was scheduled so that it ended on the final day of classes for the semester. We conjectured that students would be busier or under greater pressure at the end of the semester where they also had final exams and projects to complete. Session 3 made the task more difficult to complete by increasing the number of words to alphabetize from 150 to 200. The 1T treatments were conducted during the Spring semester of 2011 and involved 150 words. As with the 3T treatments, both deadline treatments were run at the same time.

3. Present-bias: in theory

Present-bias is typically formalized in the theoretical literature as $\beta - \delta$ quasi-hyperbolic discounting (Phelps and Pollak, 1968; Laibson, 1994, 1997 and O'Donoghue and Rabin, 1999a).[8] In this section we briefly delineate the theoretical implications of quasi-hyperbolic discounting for the simple decision problem in our experiment, so as to provide a frame of reference for the empirical results. Because we focus on quasi-hyperbolic discounting, we will often use the term interchangeably with present-bias, while subjects without present-bias may be referred to as exponential or non-present-biased. In a previous draft, Bisin and Hyndman (2018), we introduced a stopping-time problem which formally represents the decision problem in the experiment. We solved the model under both exponential and quasi-hyperbolic $\beta - \delta$ discounting, allowing for a costly effort to complete the task which changes stochastically according to a finite Markov chain. The following discussion is based on this analysis.

The decision problem is naturally divided into two distinct stages. First, in the ex-ante stage in the lab, agents choose (possibly binding intermediate) deadlines for each of the tasks they have the option to complete. The second stage is the active decision stage during which students in their free time choose whether to complete the task(s) assigned or engage in other activities. Each student's behavior depends on her discounting preferences: how patient she is and whether she is subject to a present-bias; that is whether $\beta < 1$ in $\beta - \delta$ quasi-hyperbolic discounting.

It is straightforward to show that, in this decision problem, a decision maker with exponential discounting always prefers not to self-impose any deadline since doing so only limits the option value of waiting for a lower cost. However, the same cannot be said for a present-biased decision maker. Because she knows that she may be tempted to delay in the future, she

[7] In Phase 1, students pre-address envelopes and fill in their petty-cash vouchers. This is done to both increase the credibility and saliency of payments, and to make the processing of payments easier for us.

[8] Other types of preferences may lead to procrastination and demand for commitment. Examples include the models of temptation and self-control by Gul and Pesendorfer (2001, 2004), dual-self models such as Benhabib and Bisin (2005) and Fudenberg and Levine (2006), optimal expectations and over-confidence models such as Brunnermeier et al. (2008).

may prefer to commit to an earlier deadline inducing himself/herself to complete the task on average earlier than without the deadline.[9]

While present-bias has clear and sharp implications in the (first) deadline setting stage, in the second stage it has no clear qualitative implications. Both exponential and quasi-hyperbolic students will optimally employ a threshold rule: complete the task if the cost falls below a threshold. The only difference is that the threshold for quasi-hyperbolic agents will generally be lower,[10] because present-bias introduces an incentive to procrastinate.[11]

To foreshadow our analysis, we also note that in a dynamic choice context like the one we study, besides external self-control mechanisms such as self-imposed deadlines, there are also various internal (psychological) mechanisms which subjects may employ to avoid procrastination. These internal mechanisms include mental deadlines, cues, anticipatory planning, and other self-regulatory behavior, (see e.g., Baumeister et al., 1994; Kuhl and Beckmann, 1985 and Gollwitzer and Bargh, 1996), and have been demonstrated to be activated when decision problems involve multiple repeated tasks.[12] We will return to this discussion in the context of our results in the following sections.

4. Present-bias: in the experiment

In this section, we describe several main empirical results pertaining to present-bias. First of all we report results on the students' deadline setting and task completion choices. These are the main dimensions over which present-bias is identified, as we have argued in Section 3. We document that (i) a significant subset of the students do self-impose a binding deadline, when allowed to do so; but that (ii) treatments with either exogenous or endogenous deadlines are not associated with more frequent or earlier completion. In order to more fully understand the role of deadlines in students' task completion behavior, we then identify which students are present-biased. To this end, we exploit our initial survey data to look at the psychological correlates of behavior and deadline setting. We find that (iii) present-bias is relatively widespread in our sample; that is, psychological characteristics indicative of present-bias are widespread and are correlated with setting deadlines. We also find that (iv) while present-biased students procrastinate in 1T treatments, in the repeated task treatments (3T) they successfully manage to self-control; that is, exponential and present-biased students behave very similarly.

4.1. Deadline setting

Table 2 documents that there is a *robust demand for setting deadlines* across all of our sessions with endogenous deadlines. In the One-Task treatment, 31.4% of subjects self-imposed a deadline. Conditional on setting a deadline, the average number of days before the ultimate (one-week) deadline was 1.6 days, although many of the deadlines were actually less than 12 hours before the end of the experiment.

In the Three-Task treatments, there is even greater demand for setting deadlines.[13] For the first task, depending on the session, between 33.3% and 61.9% of subjects self-imposed a binding deadline, while fewer subjects imposed binding deadlines for tasks 2 and 3, the range was still between 20.8% and 57.1%.[14] The results also suggest that subjects respond to exogenous external cost factors – timing of experiment within the semester and number of words in each task. In particular, when the task occurs at the end of the semester (when subjects should be busier) or the task is more difficult (i.e., 200 words), subjects are less likely to self-impose a deadline and the deadlines that they do impose are less strict.[15]

Broadly speaking, in the Three-Task Endogenous Deadlines treatment, there are two typical deadline profiles. Approximately 62% of deadline profiles are *strictly increasing* in that $D_1 < D_2 < D_3 \leq T$, where D_i is the deadline chosen for task i and T is the exogenously imposed end of the experiment. Of these strictly increasing deadline profiles, there was not a

[9] Theoretical studies of the effects of external commitment devices in dynamic choice environments include O'Donoghue and Rabin (1999b), who characterize general external mechanisms to induce second-best optimal behavior in agents who procrastinate due to present-bias preferences, Sáez-Martí and Sjögren (2008), who study how binding deadlines affect the timing of effort when agents get distracted, and Battaglini et al. (2005) for a theoretical analysis of commitment through peer groups.

[10] We omit any discussion of naïve quasi-hyperbolic discounters, who have a present-bias, but are unaware of it. It is very difficult to separately identify these students based on the distribution of task completions. We will provide a brief discussion of naïveté in the decision stage in Section 6.2.

[11] O'Donoghue and Rabin (1999a) show, in an example, that in their environment it is possible that quasi-hyperbolic agents preproperate (that is, that they complete the task before exponentials, ceteris paribus); see Example 3, p. 112. However, their example requires that both costs and rewards are time-varying and in special ways. This is never the case in our formulation.

[12] A large theoretical literature in psychology and economics studies the form and the effectiveness of self-control mechanisms; see e.g., Ainslie (1992, 2001); Laibson (1994). More recent work includes Benabou and Tirole (2004); Benhabib and Bisin (2005) and Hsiaw (2013).

[13] A two-sided proportions test comparing the 1T treatment with the equivalent 3T treatment, rejects the hypothesis that the proportion of subjects setting deadlines is the same ($p = 0.026$).

[14] For comparison, in studies 1 and 2 of Ariely and Wertenbroch (2002) 73% and 100% of subjects, when given the choice, self-imposed a deadline on at least one task. This could be because deadlines were soft and so there was less cost associated with missing a deadline. Augenblick et al. (2015) find that 59% of subjects demand commitment when there is no direct monetary cost.

[15] Comparing the 150 and 200 word sessions, a two-sided proportions test just misses marginal significance for this result ($p = 0.102$). Comparing the mid-semester and end-semester sessions, we can say that deadlines are significantly less strict in the end of semester implementation than in the mid-semester, 150 words implementation (conditional on setting a binding deadline: $p = 0.075$; unconditional on setting a binding deadline: $p = 0.015$).

Table 2
Self-imposed deadlines.

(a) 1T treatments Mid-semester, 150 words			
	Task 1	Task 2	Task 3
Strength of Deadline in Days (Conditional)	1.6	–	–
% of Subjects Setting Deadlines	31.4	–	–
(b) 3T treatments Mid-semester, 150 words			
	Task 1	Task 2	Task 3
Strength of Deadline in Days (Conditional)	7.7	5.6	5.1
% of Subjects Setting Deadlines	61.9	57.1	42.9
End-semester, 150 words			
	Task 1	Task 2	Task 3
Strength of Deadline in Days (Conditional)	4.5	2.5	2.3
% of Subjects Setting Deadlines	33.3	33.3	20.8
Mid-semester, 200 words			
	Task 1	Task 2	Task 3
Strength of Deadline in Days (Conditional)	5.4	4.2	3.8
% of Subjects Setting Deadlines	50.0	50.0	40.9

Table 3
Descriptive summary of the completion statistics.

	Fraction of tasks completed				
	Endogenous			Exogenous	No deadlines
	Set deadline	No deadline	Overall		
1T treatments	54.5%	58.3%	57.1%	–	45.6%
3T treatments	33.3%	40.4%	36.8%	40.6%	47.0%

Highlighted cells indicate a statistically significant difference at the 5% level or better between the two treatments (two-sided test). Note that tests are only conducted on the overall sample for the treatment and not conditional on whether or not a deadline was set in the Endogenous Deadlines treatment.

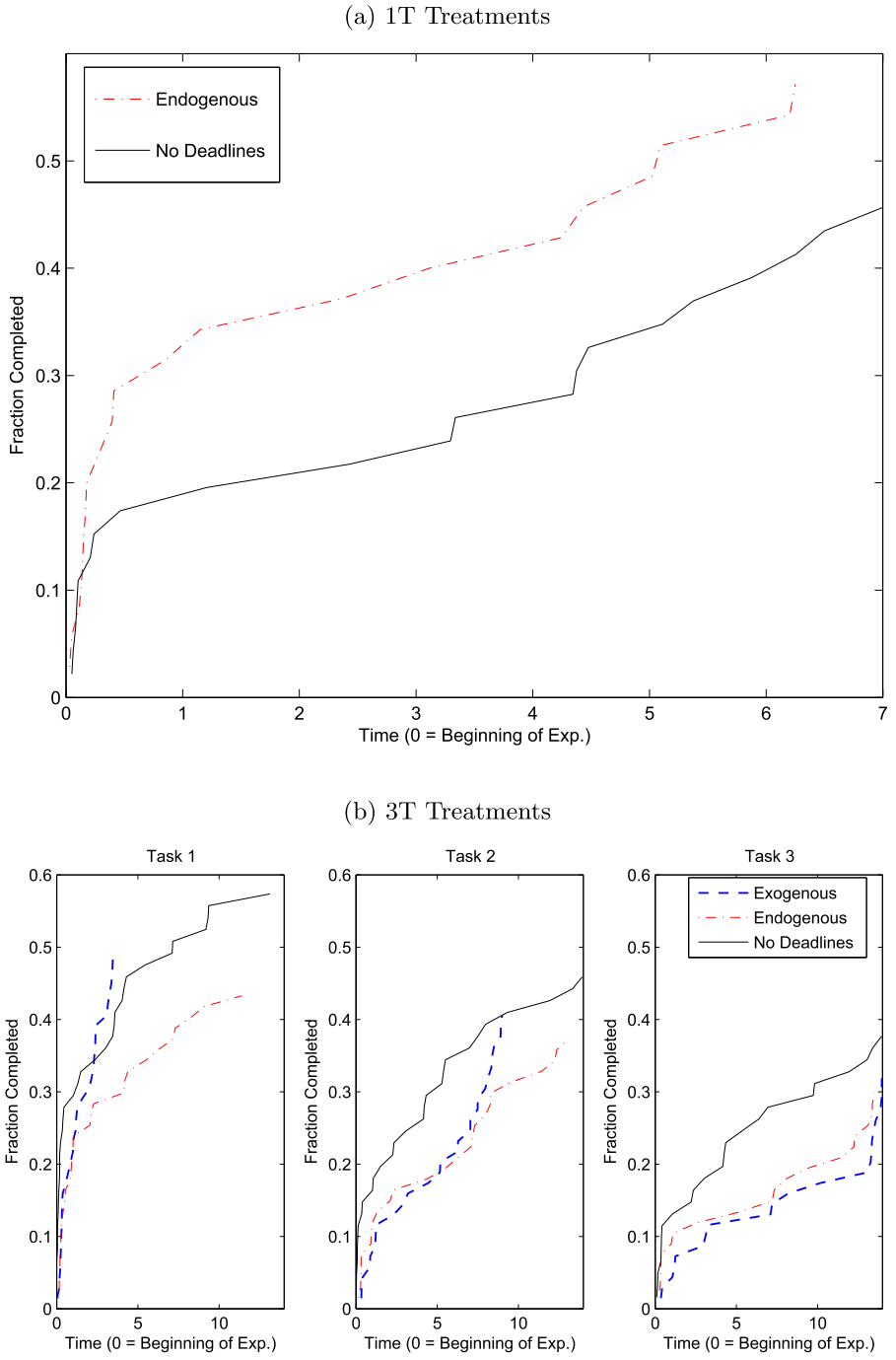
single instance in which it was close to the deadline profile in our Exogenous Deadlines treatment. Approximately one-third of deadline profiles are *(partially) flat* in the sense that $D_i = D_{i+1}$ for at least one i and $D_j < T$ for at least one j.

4.2. Task completions

Table 3 provides the overall task completion rates by experimental treatment. This is the fundamental comparison, under random assignment to treatment. Results are striking: at best, there is *no positive association between deadlines and completion*. In One-Task treatments, although the completion rate appears to be higher in the Endogenous Deadlines treatment, the difference (57.1% vs. 45.6%) is not statistically significant. In the Three-Task treatments, we actually see that the completion rate is lowest in the Endogenous Deadlines treatment (36.8%), followed by the Exogenous Deadlines treatment (40.6%) and highest in the No Deadlines treatment (47.0%). Moreover, the difference between the Endogenous and No Deadlines treatments is statistically significant at $p < 0.05$.

Analyzing the data more in detail, in Fig. 2 we show the cumulative distribution of task completions for each of our treatments and, for the Three-Task treatments, separate distributions for each task. In the One-Task treatments, the main difference is that a larger proportion of subjects completed the task almost immediately in the Endogenous Deadlines treatment. In the Three-Task treatments, comparing the Exogenous and No Deadlines treatments, the role for deadlines becomes apparent. For each task, there is a spike in completions in the time immediately preceding the deadline in the Exogenous Deadlines treatment. At least for Tasks 1 and 2, by the deadline, the completion rate meets or exceeds to completion rate in the No Deadlines treatment. However, in the No Deadlines treatment, we continue to see later task completions. Hence, by the end of the experiment, for all tasks, the completion rate is higher without a deadline.

For the Endogenous Deadlines treatments, Table 3 also reports completion rates conditional on setting or not setting binding deadlines. Interestingly, subjects who actually self-imposed a binding deadline in the Endogenous Deadlines treatments had lower completion rates than those who did not. Furthermore, there is suggestive evidence that task completions in the Endogenous Deadlines treatment differed based on the deadline profile chosen by subjects who self-imposed deadlines. In particular, subjects who imposed a strictly increasing deadline profile completed only 25.4% of tasks, while subjects who imposed a (partially) flat deadline profile completed 54.5% of tasks. This difference is weakly significant according to a Mann-Whitney rank sum test ($p = 0.0835$).



Note: The figure reports the cumulative fraction of tasks completed in each treatment over time.

Fig. 2. Cumulative distribution of task completions.

4.3. Present-biased students

According to the theory discussed in Section 3, setting a deadline is a manifestation of demand for commitment, and hence it is evidence of present-bias. However, under this interpretation, it is somewhat puzzling that performance is *lowest* under Endogenous Deadlines and *highest* under No Deadlines in the 3T treatments. In particular, given random assignment to treatment, we would have expected performance to be highest in the Endogenous Deadlines treatment. Specifically, comparing subjects in the Endogenous versus No Deadlines treatments, the performance of non-present-biased subjects should

Table 4
Self-reported characteristics and self-imposed deadlines in endogenous deadlines treatment (3T treatment, all survey questions).

		No deadline set	Deadline set	p-Value MW test
Cost Factors	# of courses	4.06	3.59	0.197
	GPA	3.42	3.45	0.509
	# of exams	0.94	0.94	0.526
	# of major assignments	1.27	1.82	0.074
	# of minor assignments	2.70	4.24	0.045
	Have job?	0.55	0.47	0.543
	# of clubs	1.09	1.68	0.077
	Time studying	29.65	30.56	0.816
	Time family	3.55	3.53	0.641
	Time job	13.80	16.34	0.990
	Time socializing	12.55	17.55	0.038
Psychological Factors	Conscientious	4.21	3.76	0.010
	Often late (assignments)	1.52	1.44	0.589
	Often on time (appointments)	4.03	4.26	0.265
	Not dependable	2.55	2.32	0.406
	Detail oriented	4.03	3.85	0.355
	Not organized	2.30	2.35	0.758
	Follow schedule	3.79	3.50	0.152
	Unexpected events	3.39	3.06	0.149
	Impatience (Ideal allocation)	5.79	6.85	0.067
	Temptation allocation	6.06	6.71	0.422
	Perceived actual allocation	5.82	6.88	0.110

Note 1: Higher numbers indicate more of the particular characteristic.
Note 2: Highlighted cells indicate that groups are (weakly) significantly different at the 10% level or better according to a Mann-Whitney rank-sum test.
Note 3: The cells dealing with impatience, temptation and perceived actual allocation are taken from Ameriks et al. (2007). The precise survey questions for all cells are given in Appendix A.2.

be identical because present-biased subjects will not set deadlines. However, present-biased subjects' performance should be better in the Endogenous Deadlines treatment because they will optimally set a deadline to improve performance.[16]

Motivated by this, we attempt to understand whether some other behavioral factor, besides present-bias, can contribute to explain deadline setting in our data. To this end, we seek to statistically identify which students are present-biased using a flexible approach which exploits deadline setting as evidence of present-bias. More specifically, we exploit the correlation between the self-reported psychological characteristics of our students we collected in the initial survey and their deadline setting behavior. This is because, as we noted in Section 3, completion data has relatively little identification power in this context (see also Heidhues and Strack, 2019; Harris and Laibson, 2013, for a formal result along these lines).

The survey results show that there are meaningful differences between those students who do and do not self-impose deadlines. Most importantly, these differences are a psychological indication of present-bias. For the 3T Endogenous Deadlines treatment, Table 4 compares subjects' responses to the various costs of time and psychological characteristics depending on whether or not they self-imposed a deadline. It shows that both psychological and cost factors influence the decision to self-impose a deadline.[17] In particular, subjects who self-impose deadlines tend to face a higher cost of their time at the margin, as indicated by having more minor and major assignments, being in more clubs and spending more time socializing (though they are in significantly fewer courses). With regards to psychological factors, subjects who self-impose a deadline also self-report to be less conscientious and more impatient.

Formally, to identify present-biased subjects we proceed by estimating a logit model using our 3T Endogenous deadlines treatment on the decision to self-impose a deadline, where the Phase 1 survey responses are the explanatory variables. Using the parameter estimates from the model, we can compute, for a subject in *any treatment*, the predicted probability that the subject would self-impose a deadline (and hence be present-biased).[18]

[16] One can make a similar argument for why performance should be higher in the Endogenous Deadlines treatment than in the Exogenous Deadlines treatment. On the other hand, a clear comparison between No Deadlines and Exogenous Deadlines is not possible. The exogenously imposed deadlines will hurt present-biased subjects, relative to no deadlines; however, they may help present-biased subjects.

[17] In testing for differences between those subjects who do and do not self-impose deadlines for each of our survey questions, it is possible that some variables will be significant purely by chance. To guard against this, we conducted a Monte Carlo exercise. Specifically, for each question and for each of 100,000 trials, we drew random sample of subjects of the same size as our Endogenous Deadlines treatment with the same mean and standard deviation as in our sample. We then tested – both with a t-test and a Mann-Whitney rank sum test – for differences between the two simulated populations; results are reported in the last two columns of Table 4. From this exercise, we can compute the probability of observing six (as we have) or more questions where we reject equality of means at the 10% level or better. For the t-test and Mann-Whitney tests, these probabilities are 0.029 and 0.027, respectively, supporting our claim that the observed significant differences are not due to chance.

[18] We estimate this logit model both using all survey questions and, more parsimoniously, by only including the most relevant variables for the decision to self-impose a deadline. This should guard against the possibility of over-fitting the model and making for worse out of sample predictions, which is our primary interest. Qualitatively, the results are not sensitive to whether the full or parsimonious model is used. Note also that we focus on the 3T

The logit analysis documents that *present-bias is relatively widespread* in our sample. As a first cut, we label a subject as "present-biased" if the predicted probability of setting a deadline is greater than 50%. According to this procedure, we classify about 42% of the subjects in our experiment as present-biased: 26 out of 81 in 1T and 90 out of 197 in 3T treatments. This is comparable with the results in Halevy (2012), which in a lab experiment conducted in class shows that 52% of subjects are not time-consistent. In lab studies using monetary payments, Casari (2009) finds that about 65% of students exhibit some form of choice reversal. Note that our model correctly classifies approximately 80% of subjects in the 3T Endogenous Deadlines treatment. Specifically, 79.4% of subjects our model labels as present-biased actually self-imposed a deadline, while 81.8% of the subjects labeled as *not* present-biased *did not* self-impose a deadline.

4.4. Present-biased behavior: procrastination and self-control

Having distinguished between exponential (non-present-biased) and present-biased subjects, based on their predicted likelihood of self-imposing a deadline, we now examine each type's distribution of task completions. In theory, as noted in Section 3, present-biased students should delay task completion with respect to exponential students. However, if the presence of repeated tasks operates as a cue to induce subjects to exercise (internal) self-control, then the difference in behavior between present-biased and exponential subjects may be less in the 3T treatments than in the 1T treatments.

Fig. 3(a) shows the distribution for subjects in the 1T treatments, while Fig. 3(b) shows the distributions (one for each task) for the subjects in the No Deadlines and Exogenous Deadlines treatments of the 3T treatments. The results are striking. In the 1T treatments, consistent with the theoretical implications of $\beta - \delta$ quasi-hyperbolic discounting, present-biased subjects show substantial delay in task completion (though, by the end of the experiment, they have an overall task completion rate that is similar to exponential subjects). In contrast, in the 3T treatments, subjects we have identified as present-biased do not delay task completion with respect to those we have identified as exponential. In other words, *present-biased students procrastinate in single-task (1T) treatments but they are very successful in exercising self-control in repeated-task (3T) treatments.* As alluded to at the end of Section 3, this behavior is consistent with the theoretical implications of the studies of self-control in repeated tasks.

4.5. Comparison with the literature

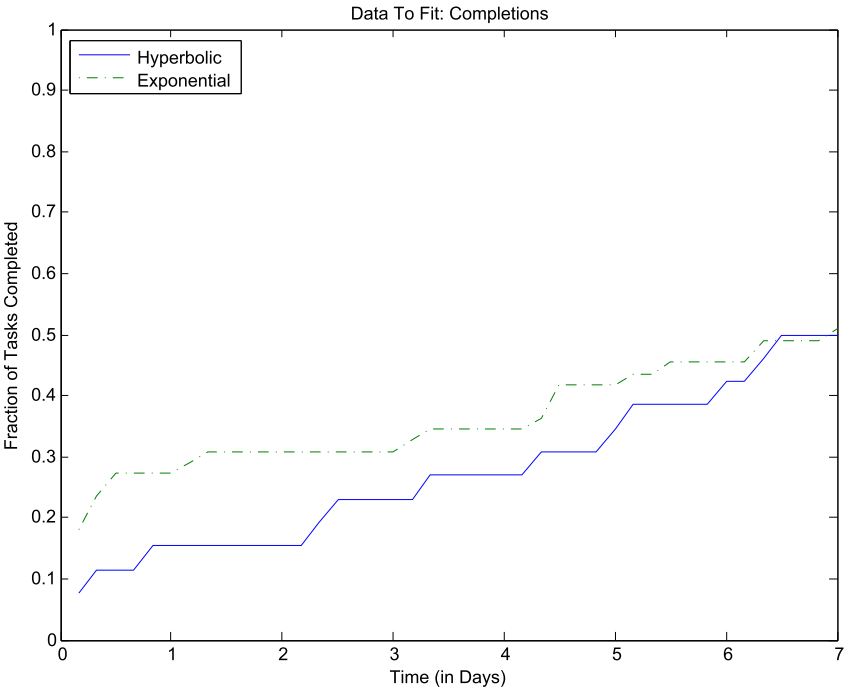
Like us, a few recent papers study procrastination in the context of students' academic work. The results are inconclusive and, in some cases, run counter to our own. Fundamentally, our data suggests that students, when given the opportunity, might not set deadlines optimally and that deadlines may turn out to be costly for students, unduly reducing their completion rates. Subjects who self-imposed a binding deadline in the Endogenous Deadlines treatments had lower completion rates than those who did not; particularly so for subjects who imposed a strictly increasing deadline profile. While it is not necessarily surprising that subjects who set deadlines had lower completion rates than those who did not (deadlines could simply be a manifestation of an awareness of present-bias, and we would expect present-biased subjects to have lower completion rates), these results appear in contrast with Ariely and Wertenbroch (2002), who reported that subjects' performance was worst in their No Deadlines treatments and that subjects in the Endogenous Deadlines treatment – especially those who set relatively evenly spaced (hence, increasing) deadlines – saw performance gains.

In fact, perhaps the contrast is less strong than it appears. Specifically, our analysis of the raw data in Ariely and Wertenbroch (2002) (kindly provided to us by the authors) shows that the gains in performance in their first study are not significant when restricted to the treatment tasks. Instead, it is the final grades (which includes the treatment tasks, a final paper, and other components) where we see performance being significantly higher in the Endogenous Deadlines treatment. Moreover, in their second study, although students' performance is higher under deadlines, they end-up disliking the task more when they are subject to deadlines, leaving some doubt about whether the effect of deadlines is effectively on procrastination.

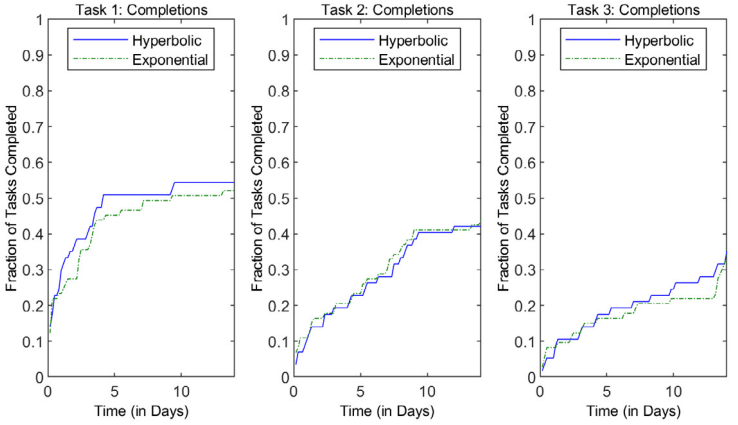
Furthermore, partially contrasting results can be explained by different experimental set-ups. Indeed, our reward mechanism differs from Ariely and Wertenbroch (2002). In our experiment, students are rewarded through a fixed, known, homogeneous monetary payment at a pre-specified delay from completion, our experiment controls for student motivation in performing tasks. This is in contrast to Ariely and Wertenbroch's Study 1 in which Executive MBA students were rewarded for (less measurable) academic performance. Second, the tasks in our experiment are the same for all students (alphabetize one (up to three) list(s) of "words") and do not require any special skill which could be heterogeneously distributed across the student pool; this is in contrast to the writing task of Ariely and Wertenbroch's Study 1 as well as to the proof-reading task of their Study 2, in which heterogeneous ability could arguably affect the results. We also impose an upper bound on the time to complete the task after initiating it, so that students are essentially required to complete each task in one sitting. Without such a restriction, as in Ariely and Wertenbroch (2002), there is no clear link between the time effort is exerted and the time the reward is obtained: students could smooth effort over time and could even trade-off effort and time, all of which makes it difficult to interpret the results of the experiments as evidence for/against procrastination due to

Endogenous Deadlines treatment, rather than the 1T Endogenous Deadlines treatment because latter is substantially less predictive of self-imposing a deadline.

(a) One Task Treatment



(b) Three Task Treatment – Excludes Endogenous Deadlines



Note: The figure reports the actual cumulative fraction of tasks completed over time for two categories of subjects: those that our logit analysis indicates are hyperbolic and those that our logit analysis indicate are exponential.

Fig. 3. Distribution of completions for exponential and quasi-hyperbolic subjects.

present-bias. Furthermore, the time restriction to complete the task we impose allows us to collect data on failed attempts, which can be exploited to better understand the determinants of students' behavior. Another distinctive feature of our design is that self-imposed deadlines are necessarily hard deadlines, while the deadlines in the Ariely and Wertenbroch (2002) experiments are "soft" in the sense that only a per-period penalty is imposed for completion after the deadline. While soft deadlines occur perhaps more naturally outside of the realm of these experiments, their theoretical implications are harder to obtain and hence it is harder to interpret any effects of such deadlines in terms of the underlying characteristics of the preferences of students which might motivate their demand for commitment and their behavior.

Our results on deadlines are more in line with those in another recent paper, Burger et al. (2011). They conduct an experiment in which students are faced with a time allocation problem over a task of significant duration (studying 75 hours over a 5-week period in a monitored location) under different constraints in the form of binding sub-deadlines (e.g., 15 hours in the first week). Unlike Ariely and Wertenbroch (2002), Burger et al. (2011) do not allow subjects to set

binding deadlines. The main result of the paper is that deadlines do not lead to more students successfully completing the task. Indeed, the experimental design adopted by Burger et al. (2011) is more similar to ours in the sense that student behavior – in their paper, time spent in the study room – is also unaffected by possibly heterogeneous skills and is clearly measurable; also, the monetary reward mechanism is clearly specified and so is the delay with respect to completion at which it is obtained.

5. Deadline setting behavior: partial naïveté and lack of perseverance

Our data suggest that present-bias is widespread among the students in the experiment, but that the behavior it induces is different depending on the nature of the task(s): with a single task, present-biased subjects procrastinate, while with repeated tasks, as in 3T, they appear able to successfully self-control. According to theory, present-bias manifests itself as a willingness to self-impose binding deadlines – something which we also observed in our Endogenous Deadlines treatments. However, interpreting the large demand for deadlines we observe in the experiment simply as due to present-bias might not be completely correct. First of all, as we noted, while 80% of subjects identified as present-biased set deadlines when given the chance, about 20% do not. This is consistent with the theory in Section 3; as we show in Bisin and Hyndman (2018), depending on the underlying cost and present-bias parameters, present-biased subjects may not set deadlines. It is somewhat more problematic to explain the 18% of non-present-biased subject who self-impose deadlines as well as the fact that present-biased students self-impose binding deadlines even though they are able exercise self-control in the 3T No Deadlines treatment. Indeed, in the case of repeated-task treatment, if students *correctly* anticipated their ability to self-control, then they should not have self-imposed deadlines. This is because deadlines are just as costly for a decision maker who can perfectly self-control as they are for a non-present-biased decision maker.

In this section, we examine the behavioral components of students' anticipation formation mechanisms. We argue that, together with present-bias, these mechanisms can rationalize our data. More specifically, we argue that our data suggests two kind of complementary behavioral anticipation formation mechanisms.

1. Students under-estimate their ability to exercise self-control (at least in repeated tasks), a form of *partial naïveté at the deadline setting stage*.
2. Students are over-confident about their ability to complete the task; in particular they over-estimate their ability to persevere on the current and future tasks after a failed attempt.

These mechanisms are complementary in the sense that both might lead students believe that deadlines are more beneficial than they actually are, leading them to self-impose deadlines, which ends up harming them.

Fundamentally, deadlines have three distinct effects: (i) they provide commitment; (ii) they impose a cost in terms of the option value of waiting for a lower cost of task completion; (iii) they amplify the cost associated with over-confidence and lack of perseverance. We argue that our data suggest that partial naïveté and over-confidence might have a negative effect on present-biased students who are able to self-control (and possibly exponential students who might not fully realize they are not present-biased). This would be especially the case in repeated-task experiments, where self-control is widespread in our data. Indeed, as Table 3 showed, students who set deadlines in 3T complete fewer tasks.

5.1. Partial naïveté in optimal deadline choice

Many of the subjects we have identified as quasi-hyperbolic present-biased (and a few of the exponential discounters) choose to self-impose stringent deadlines in the 3T treatments: the 95% confidence interval on the fraction of self-imposed deadlines on at least one task is (38.24%, 63.18%). On the other hand, Fig. 3(b) showed that quasi-hyperbolic subjects' completion behavior was virtually indistinguishable from exponential subjects. As we noted, this is puzzling but can be rationalized if subjects are partially naïve about their ability to internally self-control.

Recall that the choice of deadlines in our experiment is made at the ex-ante stage, in the lab. It is natural to distinguish conceptually between two different present-bias parameters: one is operational at the ex-ante (deadline choice) stage in the lab and the other is operational in the active decision stage in their daily life outside the lab. Even if, as we assume, quasi-hyperbolic agents are sophisticated in the decision stage, they may not correctly anticipate their ability to exert internal self-control while in the lab. The same might hold for some of the exponential discounters. In this case, we say that subjects are *partially naïve*.[19] The main behavioral implication for such subjects is that they may make sub-optimal deadline choices.

[19] Evidence for partial naïveté in dynamic choice contexts has been suggested by DellaVigna and Malmendier (2006) in their study of gym memberships and attendance. More recently Augenblick and Rabin (2019), study subjects who choose how much of an unpleasant task to complete immediately for various payment schemes. They find strong evidence for present bias, but also that subjects only anticipate 10-24% of their present bias. Fang and Wang (2015) provide a method for estimating dynamic discrete choice models and also find evidence for partial naïveté in women's decisions to undergo mammograms. More closely related is the field study by John (2017). She shows that partially naïve subjects are more likely to choose weak commitment devices which lead to eventual default. We show that partial naïveté about one's ability to internally self-control, leads subjects to self-impose binding deadlines in the multiple task treatment, to their detriment.

Table 5
Classification of effort.

(a) Three task, endogenous

Task	No attempt	Give up	finish
1	0.239	0.328	0.433
2	0.522	0.104	0.373
3	0.642	0.060	0.299
Tot	0.468	0.164	0.368

(b) Three task, exogenous

Task	No attempt	Give up	Finish
1	0.203	0.304	0.493
2	0.449	0.145	0.406
3	0.623	0.058	0.319
Tot	0.425	0.169	0.406

(c) Three task, no deadlines

Task	No attempt	Give up	Finish
1	0.115	0.311	0.574
2	0.475	0.066	0.459
3	0.574	0.049	0.377
Tot	0.388	0.142	0.470

(d) One task

Treat	No attempt	Give up	Finish
N.D.	0.130	0.413	0.457
Endog	0.229	0.200	0.571

Note: The variable "No Attempt" reports the fraction of subjects who never submitted a list of words for the task in question; "Give Up" reports the fraction of subjects who submitted at least one list of words for the task but ultimately did not successfully complete the task; "Finish" reports the fraction of subjects who successfully completed the task in question.

In the 1T treatment, it is difficult to directly say whether and how subjects misperceive their ability to self-control. In Endogenous Deadlines treatment the 95% confidence interval on the fraction of self-imposed deadlines is $(16.85\%, 49.29\%)$, suggesting that subjects anticipate an inability to self-control. Fig. 3(a) shows that present-biased subjects procrastinate. However, recalling Table 3 for the Endogenous Deadlines shows that there is no appreciable difference in completion rates between subjects who did and did not self-impose deadlines ($p = 0.834$). This suggests that subjects may be approximately correctly calibrated about their ability to self-control.

5.2. Overconfidence and lack of perseverance in task completion

Our data does not contain just successful completions, but also each attempt to complete the task(s) that subjects made. From these data, we document that students appear to under-estimate the likelihood of making mistakes in completing the tasks and over-estimate their ability to persevere through these mistakes. To see this, first observe that in the 3T treatments, subjects reported a belief of between 83 and 90% that they would finish all three tasks. In 1T, beliefs about completing the task are approximately 86% (Endogenous) and 91% (No Deadlines). Next, note that for each (subject, task) pair, one of three things can happen: (i) either the subject makes no attempt on the task, (ii) the subject makes an attempt on the task but ultimately gives up working on the task without successfully completing it or (iii) the subject successfully completes the task.[20] Table 5, reports the frequency of each possible outcome for each treatment and task.

Across all three deadline conditions in the 3T treatments, we see that between 14.2 and 16.9% of tasks are attempted but not completed. That is, the subject gives up. These numbers account for at least part subjects' apparent over-confidence (i.e., the gap between subject's beliefs and their actual completion behavior). This suggests to us that subjects either do not expect to make mistakes of that they expect to be able to persevere. In fact, we would argue that the impact of a lack of perseverance is even larger than the direct effect of giving up. As can also be seen from Table 5, the frequency of "No Attempt" is increasing in task number, which indicates that subjects may be dropping out (i.e., quitting) the experiment. Indeed, Table 6, which shows the complete pattern of transitions (pooled over deadline treatments because the results are similar across deadline treatments) shows that such drop-out behavior in later tasks is driven by giving up on a previous task. For example, conditional on giving up task 1, 91.94% of subjects make no attempt on task 2. On the other hand, of the subjects who finish task 1, 79.59% go on to finish task 2.

Tables 5 and 6 additionally point to the hidden cost of deadlines. As can be seen from Table 5, the frequency of "No Attempt" and "Give Up" are almost always higher in the Exogenous and Endogenous Deadlines treatments than in the No Deadlines treatments. As we now show, deadlines play an important role in explaining this result. To see this, we examine how the likelihood of success varies with the time a subject has left (before the deadline) when they begin working on a task. Fig. B.1(a) in the appendix shows that, if anything, subjects who face binding intermediate deadlines start working earlier in absolute terms than subjects who face no deadlines. However, Fig. B.1(b) shows that there is much less time available until the deadline. Finally, Fig. B.1(c) shows that subjects who begin working on a task closer to the deadline are much less likely to ultimately complete the task. Specifically, subjects who begin a few hours before the deadline have a less-than 50% chance of completing the task, while subjects who begin a week or more before the deadline have a 70% chance or higher of completing the task. Indeed, we draw the same conclusion as in Fig. B.1(c) with a random effects logit regression. The estimated coefficient on time remaining is positive and significant ($p = 0.002$). For every additional day

[20] Note that we consider an attempt to be the first time a subject submits a list of words for each task issued. For the same issued task, it may take several submissions before the task is successfully completed, but in our analysis, these are not considered separate attempts. Recall that a new task is issued each time a subject logs in, refreshes the screen or submits a task after 2 hours since it was issued.

Table 6
Transitions of effort type from task to task.

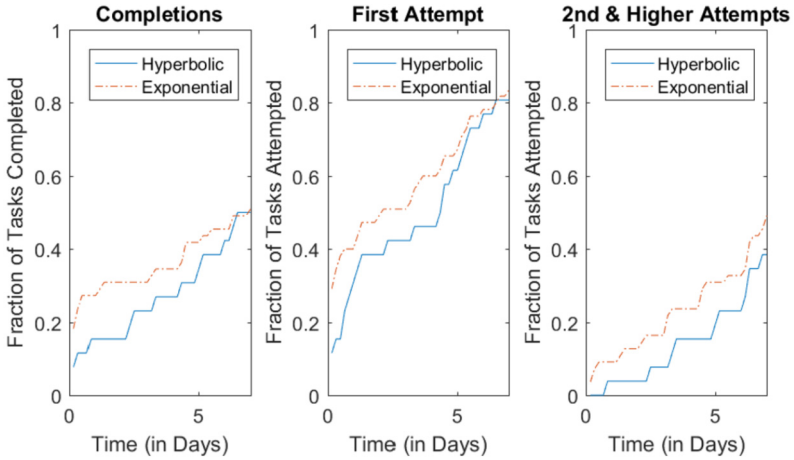
(a) Task 2 effort given outcome of Task 1

		Task 2		
		No att.	Give up	Finish
Task 1	No Att.	81.08	13.51	5.41
	Give Up	91.94	6.45	1.61
	Finish	8.16	12.24	79.59

(b) Task 3 effort given outcome of Task 2

		Task 3		
		No att.	Give up	Finish
Task 2	No Att.	96.84	2.11	1.05
	Give Up	85.71	0.00	14.29
	Finish	13.58	11.11	75.31

Note: The table should be read as follows: Each cell represents the probability of outcome along the column in task $n+1$, conditional upon the outcome along the column in task n, $n = 1, 2$. For example, in the upper-left cell of panel (a), it means that 81.08% of subjects who made no attempt on task 1 also made no attempt on task 2.



Note: The figure reports, the actual cumulative fraction of tasks completed or attempted over time for two categories of subjects: those that our logit analysis indicates are hyperbolic and those that our logit analysis indicate are exponential. The middle panel focuses on the timing of a subject's first attempt only, while the right-most panel looks at second and higher attempts.

Fig. 4. The cumulative fraction of tasks completed & attempted (1T treatments).

before the deadline that one starts a task, the probability of completing the task increases by approximately 2.6%. Thus, because deadlines reduce the time available to find and correct mistakes, subjects who face deadlines are *less* likely to successfully complete a task that they attempt. To complete the circle, Table 6 shows that subjects who give up on one task are very likely to make no attempt on later tasks.

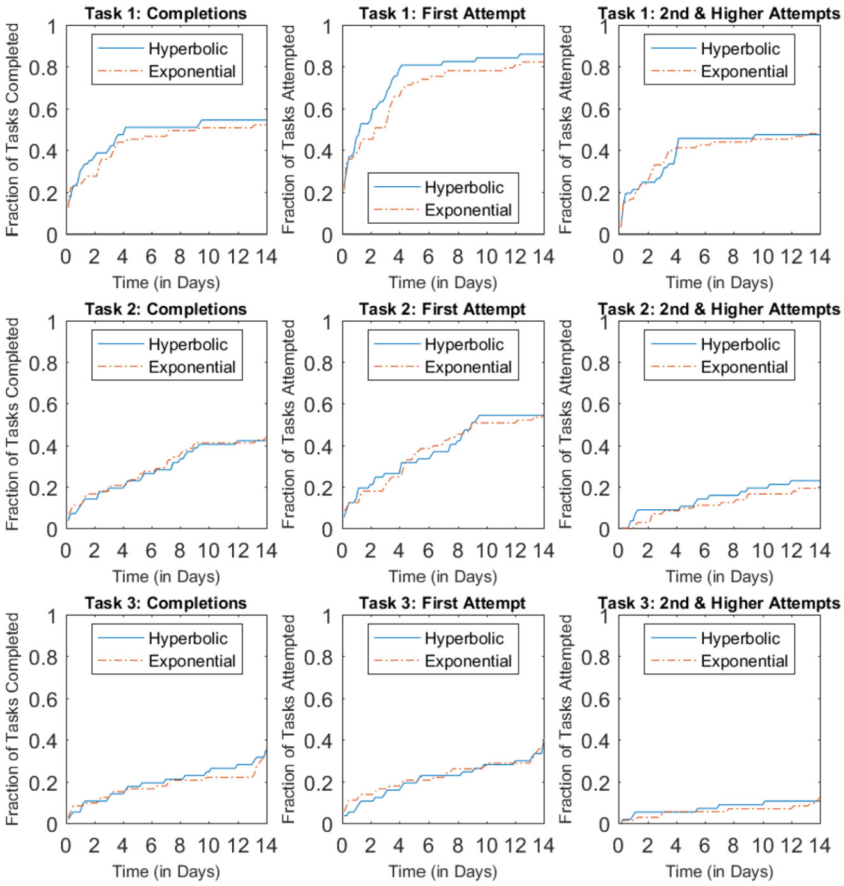
Note that the lack of perseverance we have documented has negative effects on all students, not just present-biased students. On the other hand, the costs of over-confidence and lack of perseverance hit present-biased students more. This is because the costs of lack of perseverance act as a reduction in the option value to wait for a lower cost of completing the task and hence they are amplified by deadlines, which in turn are more frequently self-imposed by present-biased students.

5.2.1. Psychological correlates of lack of perseverance

An interesting question is whether lack of perseverance and present-bias represent similar or distinct phenomena (in the sense that lack of perseverance affects both exponential and quasi-hyperbolic subjects)? Our results suggest that they may be relatively distinct. For the 1T treatments, Fig. 4 plots the observed distributions of task completions, first attempts as well as second and higher attempts for both quasi-hyperbolic and exponential discounters according to our logit classification of subject types. Just as was the case for completions (the first panel), quasi-hyperbolic discounters delay their attempts relative to exponential discounters, though for completions and first attempts these differences disappear by the final deadline. Only for second and higher attempts do quasi-hyperbolic subjects have a modestly lower rate by the end of the experiment.

For the 3T treatments, Fig. 5 plots, for each of the three tasks, the observed distributions of task completions, first attempts as well as second and higher attempts for both quasi-hyperbolic and exponential discounters. Just as with completions, the differences between exponential and quasi-hyperbolic discounters for attempts are very small, with quasi-hyperbolic discounters delaying somewhat less. Thus for both the 1T and 3T treatments, the distributions of attempts for quasi-hyperbolic and exponential subjects have the same qualitative features as the distributions of completions. This is the first piece of evidence that lack of perseverance and present-bias are distinct.

To see more evidence, note that both exponential and quasi-hyperbolic subjects have approximately equal success rates: Of those subjects who have at least one attempt, 36.96% of exponential and 38.1% of quasi-hyperbolic discounters succeed on first attempt. Of those who are unsuccessful on their first attempt, 62.1% of exponential and 61.5% of quasi-hyperbolic



Note: The figure reports, the actual cumulative fraction of tasks completed or attempted over time for two categories of subjects: those that our logit analysis indicates are hyperbolic and those that our logit analysis indicate are exponential. The middle panels of each row focus on the timing of a subject's first attempt only, while the right-most panels of each row look at second and higher attempts.

Fig. 5. The cumulative fraction of tasks completed & attempted (3T treatments).

discounters do not make any further attempts. Finally, conditional on at least one failed attempt, exponential discounters have an average of 1.93 attempts and quasi-hyperbolic discounters an average of 1.77 attempts. None of these differences are statistically significant.[21]

Finally, our survey data also suggest that lack of perseverance is a distinct psychological trait. To show this, we compare those subjects who were successful at the experiment with those who gave up at least once — indicating a lack of perseverance. For the Three-Task treatments, Table 7 reports the average response to each of the Phase 1 survey questions for subjects who either completed all three tasks or who gave-up on at least one task.

It is interesting that cost factors seem to play very little role in differentiating between subjects who complete all three tasks and those who give up at least once. There is some suggestion that subjects who complete all tasks have a higher GPA, and they have fewer exams over the course of the experiment. Instead, it appears that psychological factors drive differences. Comparing the two groups of subjects, we see that subjects who give-up at least once are significantly different for 7 of 11 psychological factors. For all 7 questions where differences arise, subjects who complete all three tasks are "better" than the subjects who give up. It is also important to note that there is no overlap in the determinants of setting a deadline (Table 4) and the determinants of giving up versus completing all tasks (Table 7).[22] This further reinforces that lack of perseverance is distinct from present-bias.

[21] More generally, subjective uncertainty might play an interesting role in choice over time; see e.g., Chatterjee et al. (2018). In our context, subjective uncertainty could be about ability to complete the task, both in terms of self-control and in terms of technical difficulty of the task itself.

[22] For the One-Task treatments (which we relegate to Table B.1 in the appendix), none of the psychological factors differ based on whether not the subject completed the task, while some cost factors do play a role (number of courses and number of minor assignments). Thus, it is fair to say that the drivers of task completion are different depending on whether subjects have one task or three tasks. In the latter case, psychological factors seem to play a greater role in determining task completion.

Table 7
Self-reported characteristics and subjects who give-up versus subjects who complete all three tasks (3T treatment).

		Give up at least once	Complete 3 tasks	p-Val. (MW test)
Cost Factors	# of courses	4.26	4.05	0.603
	GPA	3.40	3.49	0.080
	# of exams	1.37	1.00	0.051
	# of major assignments	1.68	1.45	0.310
	# of minor assignments	3.78	3.30	0.405
	Have job?	0.44	0.55	0.207
	# of clubs	1.47	1.23	0.318
	Time studying	33.76	31.53	0.421
	Time family	4.03	4.30	0.561
	Time job	10.96	12.53	0.389
	Time socializing	14.15	16.51	0.337
Psychological Factors	Conscientious	4.06	4.031	0.916
	Often late (assignments)	1.56	1.33	0.019
	Often on time (appointments)	4.03	4.50	0.003
	Not dependable	2.58	2.20	0.024
	Detail oriented	3.88	4.17	0.019
	Not organized	2.48	1.90	0.000
	Follow schedule	3.58	3.95	0.030
	Unexpected events	3.32	3.03	0.052
	Impatience (Ideal allocation)	6.11	5.92	0.992
	Temptation allocation	6.58	6.25	0.418
	Perceived actual allocation	6.47	6.12	0.547

Note 1: Higher numbers indicate more of the particular characteristic.
Note 2: Highlighted cells indicate that groups are (weakly) significantly different at the 10% level or better according to a Mann-Whitney rank-sum test.
Note 3: The precise survey questions are given in Appendix A.2.

6. Discussion

In this section we discuss several alternative behavioral explanations for some of our experimental data.

6.1. Psychological costs of self-control

Although we find the partial naïveté explanation compelling, there may be other possible alternative explanations for the observed deadline choice behavior. The fact that many of the subjects we have identified as quasi-hyperbolic discounters choose to self-impose stringent deadlines in the 3T treatments, even though they behave similarly to exponential discounters in the No Deadlines treatments, could be explained with self-control preferences along the lines of those formulated by Gul and Pesendorfer (2001, 2004), rather than $\beta - \delta$ present-bias preferences. Subjects with *Gul-Pesendorfer* preferences can be represented as facing psychological costs to exercising internal self-control mechanisms. It is possible that an external commitment mechanism (i.e., binding deadlines) might be a substitute for their own internal (and psychologically costly) mechanisms, all-the-while successfully exercising self-control in the stopping-time problem. Thus self-imposing deadlines (even those that appear not to help) could be an optimal decision because they reduce the cost of exercising internal self-control. However, it is not clear if or how such substitution works. For example, as noted in Section 6.3, in a follow-up experiment, many subjects declined to receive reminders – an alternative external mechanism to deadlines – even at zero cost and instead chose to impose binding deadlines. How to distinguish the substitution of external and internal self-control mechanisms from the commitment demand of partially naïve subjects is an open question that deserves further study. Indeed, Toussaert (2018) has some interesting work identifying *Gul-Pesendorfer* self-control types.

Another possible explanation of our experimental deadline data is that in 3T treatments we select relatively poor planners rather than present-biased agents into demanding for deadlines.[23] This interpretation is consistent with recent results showing that the nature of commitment devices alters selection (Beshears et al., 2015). It is also, in principle, consistent with the remark in Footnote 18 suggesting that commitment demand (deadlines) in 3T and 1T might represent somewhat different phenomena. However, if it were the case that commitment in 3T selects subjects who have problems planning their time rather than present-biased subjects, we would expect these subjects to perform significantly better when they can self-impose deadlines than when they cannot. We do not observe this, even when comparing subjects we have identified as quasi-hyperbolic in the different treatments. In other words, while we cannot completely exclude selection of poor planners in 3T, we believe the data do not clearly support this interpretation and we leave identifying poor planners from present-biased subjects for future study.

[23] We thank an anonymous referee for this suggestion.

6.2. Naïveté in the decision stage

The behavioral literature on present-bias distinguishes between two classes of quasi-hyperbolic discounters: sophisticated and naïve, depending on whether or not they are aware of their present-bias. In our theoretical and empirical analysis we allow for *partial naïveté* at the ex-ante stage, but do not in the stopping-time problem, effectively classifying all quasi-hyperbolic discounters as sophisticated. We now provide a discussion of this issue.

First, our procedure to identify quasi-hyperbolic discounters is really designed to identify sophisticated quasi-hyperbolic discounters. This is because it is based on the self-reported characteristics of those who self-impose binding deadlines. A naïve quasi-hyperbolic discounter – who is unaware of her present-bias – would never self-impose binding deadlines. In Table 4 we provided summary statistics on our survey questions, differentiating between those who did and did not self-impose deadlines in the 3T Endogenous deadlines treatment. As we saw, the most significant difference between students who do and do not self-impose deadlines is how they answer the conscientiousness question. Specifically, those who *do* self-impose deadlines report themselves to be *less* conscientious. This is supportive of the notion that it is the sophisticated students who are willing to impose a deadline on themselves. Thus we feel that the group we label as quasi-hyperbolic can confidently be assumed to be composed of sophisticated quasi-hyperbolic students.

It is possible that some naïve quasi-hyperbolic discounters remain hidden in the group of students we classify as exponential discounters. It might be argued that this explains why we find no difference between exponential and quasi-hyperbolic subjects in the 3T treatment. The combination of naïve quasi-hyperbolic and exponential subjects might lead to behavior which is indistinguishable from sophisticated quasi-hyperbolic subjects. However, if this were the case, we would expect the same to happen in 1T as well. Yet, as we have shown, quasi-hyperbolic subjects delay more than exponential subjects in the one task treatments. Thus, our results are in partial accordance with Mahajan and Tarozzi (2011) who shows that even if naïve quasi-hyperbolic subjects make up a substantial portion of the total population of quasi-hyperbolic subjects, they display a much smaller present bias.

In an attempt to dig more deeply regarding the identification of naïve quasi-hyperbolic discounters, we can exploit a series of questions we asked in the pre-experimental survey. These questions were previously used by Ameriks et al. (2007) to gauge this same issue. Specifically, students are asked to consider being given 10 restaurant vouchers that were valid for two years at *any* restaurant and are then asked the following hypotheticals:

q14 From your current perspective how many vouchers would you like to use in year 1?
q16 If you were to give in to your temptation, how many vouchers do you think you would use in year 1?
q17 Based on your most accurate forecast of how you think you would actually behave, how many vouchers would you use in year 1?

Following Ameriks et al. (2007), we define $SCP = q17 - q14$. This variable can be adopted as a measure of self-control problems, and in particular $SCP > 0$ can be interpreted as evidence of present-bias. However, since these questions refer to a specific context of self-control (spending on restaurants), a subject could be unaware of their general self-control problems/present-bias (i.e., they could be naïve) while still eliciting $SCP > 0$. Let the variable $SOPH$ take value 1 if our logit analysis indicates the subject is sophisticated and 0 otherwise. We can then identify naïve quasi-hyperbolic students as those who, we identify as not sophisticated (i.e., $SOPH = 0$) but who, by their survey responses, indicate self-control problems (i.e., $SCP > 0$).[24] This is the case for about 20.8% of students in the 3T treatments. On average, these students finish 1.195 tasks, while the non-naïve finish 1.25 tasks. The difference is not significant. Once more this analysis is consistent with the interpretation that, even if naïfs are present, they display a small present-bias as in Mahajan and Tarozzi (2011).

6.3. Lack of attention

Recent work by Altmann et al. (2018), Ericson (2017), Taubinsky and Rees-Jones (2018) and Taubinsky (2014) has suggested that there may be an interaction between attention (or forgetfulness) and procrastination, which is not driven by present-bias. Given that the 3T treatments take place over a two-week period, it is possible that students who fail to complete some tasks simply forget about the experiment. To test this, at the end of Spring 2011, we ran a fourth session with three treatments. In all three treatments, students are able to set their own deadlines and each task consists of 200 words. The first treatment is a baseline where no reminders are possible. In the second treatment, students can request, at no cost, to receive a reminder. In the third treatment, students can request to receive a reminder at a cost of \$3 (deducted from the participation fee). Reminders are sent via email daily at approximately 9:00AM and they inform the student of her deadlines and also provide the URL to the experimental software. We draw two conclusions from the data we obtain in these sessions.[25] First, no student is willing to pay \$3 out of her participation fee in order to receive a daily reminder. Moreover, even

[24] In support of this analysis, it turns out that there is a significantly negative relationship between the predicted probability of setting a deadline and SCP. That is, as SCP increases, we are less likely to label that subject as sophisticated. Moreover, we cannot reject that the correlation between SCP and one's self-reported level of conscientiousness (which was a key factor in the decision to self-impose a deadline) is 0 ($p = 0.45$).

[25] We only discuss the lab data, the set-up of reminders and deadlines, not the behavioral data because a glitch in the software corrupted the latter.

when reminders are free, 25% of the students choose to not receive them. Second, the presence of reminders makes people more likely to self-impose a deadline. Specifically, in the absence of reminders, 6 out of 16 subjects self-imposed a deadline on at least one task, which is comparable to the 9 out of 24 who did the same in the previously reported end of semester session. In contrast, when subjects had the option to receive reminders 31 out of 39 subjects chose to self-impose a deadline on at least one task. A proportions test easily rejects equality of proportions ($p < 0.01$). If subjects used deadlines as *de facto* reminders, then giving subjects the ability to receive reminders should lead to *fewer* self-imposed deadlines. Instead, we saw the opposite.[26] We also note that, when reminders are possible, there is no difference in the frequency with which subjects set deadlines, given whether they requested or declined to receive the reminders. Hence, it is not the case that subjects are more willing to set deadlines *because* the presence of reminders helps them remember the deadline. This confirms our prior interpretation that binding self-imposed deadlines are a manifestation of students' demand for commitment. Therefore, we do not find strong reasons to suggest that our estimates of present bias are affected by limited attention.

7. Conclusions

In this paper we study procrastination in the context of a field experiment. We find that students display a substantial demand for commitment in the form of self-imposed deadlines. However, deadlines do not appear to increase task completion rates. Present-bias, and hence procrastination, appear to constitute an important determinant of students' behavior in the experiment. Importantly, however, the effects of present-bias appear to be un-done by internal self-control when subjects engage in repeated similar tasks, as in our 3T treatments.

The behavior of students when setting deadlines remains somewhat puzzling if interpreted simply as induced by present-bias. In contrast with some evidence in the literature, notably Ariely and Wertenbroch (2002) we find that deadlines, even when self-imposed by the students, are not associated to higher completion rates and are possibly costly for the students. We argue that to understand why it is important to examine the behavioral components of students' anticipation formation mechanisms. In this respect we argue that our data suggests two kind of complementary behavioral anticipation formation mechanisms: *partial naïveté at the deadline setting stage* and over-confidence about the ability to complete the task and to persevere on a task after failing to complete it. This is consistent with the analysis of Gabaix and Laibson (2017) who show that seemingly present-biased behavior may actually arise due to misperceptions. We admit however that more work is necessary to distinguish the demand for commitment of partially naïve subjects from that of subjects trading-off the costs of internal and external self-control mechanisms.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geb.2019.11.010.

References

Ainslie, G., 1992. Picoeconomics. Cambridge University Press, Cambridge, UK.
Ainslie, G., 2001. Breakdown of Will. Cambridge University Press, Cambridge, UK.
Altmann, S., Traxler, C., Weinschenk, P., 2018. Deadlines and Cognitive Limitations. IZA Discussion Paper No. 11129.
Ameriks, J., Caplin, A., Leahy, J., Tyler, T., 2007. Measuring self-control problems. Am. Econ. Rev. 97 (3), 966–972.
Ariely, D., Wertenbroch, K., 2002. Procrastination, deadlines, and performance: self-control by precommitment. Psychol. Sci. 13 (3), 219–224.
Ashraf, N., Karlan, D., Yin, W., 2006. Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. Q. J. Econ. 121 (2), 635–672.
Augenblick, M., Rabin, N., 2019. An experiment on time preference and misprediction in unpleasant tasks. Rev. Econ. Stud. 86 (3), 941–975.
Augenblick, N., Niederle, M., Sprenger, C., 2015. Working over time: dynamic inconsistency in real effort tasks. Q. J. Econ. 130 (3), 1067–1115.
Battaglini, M., Benabou, R., Tirole, J., 2005. Self-control in peer groups. J. Econ. Theory 112, 848–887.
Baumeister, R.F., Heatherton, T.F., Tice, D.M., 1994. Losing Control: How and Why People Fail at Self Regulation. Academic Press, San Diego.
Benabou, R., Tirole, J., 2004. Willpower and personal rules. J. Polit. Econ. 112 (4), 848–886.
Benhabib, J., Bisin, A., 2005. Modeling internal commitment mechanisms and self-control: a neuroeconomics approach to consumption–saving decisions. Games Econ. Behav. 52, 460–492.
Beshears, J., Choi, J.J., Harris, C., Laibson, D., Madrian, B.C., Sakong, J., 2015. Self Control and Commitment: Can Decreasing the Liquidity of a Savings Account Increase Deposits? NBER Working Paper 21474.
Bisin, A., Hyndman, K., 2018. Present-Bias, Procrastination and Deadlines in a Field Experiment. NBER Working Paper 19874.
Brunnermeier, M., Papakonstantinou, F., Parker, J., 2008. An Economic Model of the Planning Fallacy. Working Paper.
Bryan, G., Karlan, D., Nelson, S., 2010. Commitment devices. Annu. Rev. Econ. 2, 671–698.
Burger, N., Charness, G., Lynham, J., 2011. Field and online experiments on self-control. J. Econ. Behav. Organ. 77, 393–404.
Cadena, X., Schoar, A., Cristea, A., Delgado-Medrano, H.M., 2011. Fighting Procrastination in the Workplace: An Experiment. NBER Working Paper No. 16944.
Casari, M., 2009. Pre-commitment and flexibility in a time decision experiment. J. Risk Uncertain. 38 (2), 117–141.
Chatterjee, K., Krishna, R.V., Sopher, B., 2018. Intertemporal Planning with Subjective Uncertainty, or Anticipating Your Lazy, Disorganized Self, Rutgers. Mimeo.
DellaVigna, S., Malmendier, U., 2006. Paying not to go to the gym. Am. Econ. Rev. 96, 694–719.
Duflo, E., Kremer, M., Robinson, J., 2011. Nudging farmers to use fertilizer: evidence from Kenya. Am. Econ. Rev. 101 (6), 2350–2390.

[26] See Cadena et al. (2011) and Himmler et al. (2019) for somewhat different results on the relationship between reminders and procrastination.

Ericson, K.M., 2017. On the interaction of memory and procrastination: implications for reminders, deadlines and empirical estimation. J. Eur. Econ. Assoc. 15, 692–719.
Fang, H., Wang, Y., 2015. Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. Int. Econ. Rev. 56 (2), 565–596.
Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. J. Econ. Lit. 40, 351–401.
Fudenberg, D., Levine, D.K., 2006. A dual-self model of impulse control. Am. Econ. Rev. 96 (5), 1449–1476.
Gabaix, X., Laibson, D., 2017. Myopia and Discounting. NBER Working Paper #23254.
Giné, X., Karlan, D., Zinman, J., 2010. Put your money where your butt is: a commitment contract for smoking cessation. Am. Econ. J. Appl. Econ. 2 (4), 213–235.
Gollwitzer, P.M., Bargh, J.A., 1996. The Psychology of Action. Guilford, New York.
Gul, F., Pesendorfer, W., 2001. Temptation and self-control. Econometrica 69, 1403–1436.
Gul, F., Pesendorfer, W., 2004. Self-control and the theory of consumption. Econometrica 72, 119–158.
Halevy, Y., 2012. Time Consistency: Stationarity and Time Invariance. Working Paper.
Harris, C., Laibson, D., 2013. Instantaneous gratification. Q. J. Econ. 128 (1), 205–248.
Heidhues, P., Strack, P., 2019. Identifying Present-Bias from the Timing of Choices. SSRN Working Paper 3386017.
Himmler, O., Jäckle, R., Weinschenk, P., 2019. Soft commitments, reminders, and academic performance. Am. Econ. J. Appl. Econ. 11 (2), 114–142.
Houser, D., Schunk, D., Winter, J., Xiao, E., 2010. Temptation and Commitment in the Laboratory. Working Paper.
Hsiaw, A., 2013. Goal-setting and self-control. J. Econ. Theory 148 (2), 601–626.
John, A., 2017. When Commitment Fails – Evidence from a Field Experiment. Working Paper.
Kaur, S., Kremer, M., Mullainathan, S., 2010. Self-control and the development of work arrangements. Am. Econ. Rev. Pap. Proc. 100, 624–628.
Kuhl, J., Beckmann, J. (Eds.), 1985. Action Control: From Cognition to Behavior. Springer-Verlag, Berlin.
Laibson, D., 1994. Essays in Hyperbolic Discounting. Ph.D. thesis. Massachusetts Institute of Technology.
Laibson, D., 1997. Golden eggs and hyperbolic discounting. Q. J. Econ. 113, 443–477.
Loewenstein, G., Prelec, D., 1992. Anomalies in intertemporal choice: evidence and interpretation. Q. J. Econ. 107 (2), 573–597.
Mahajan, A., Tarozzi, A., 2011. Time Inconsistency, Expectations and Technology Adoption: The Case of Insecticide Treated Nets. Stanford University. Mimeographed.
Novarese, M., Giovinazzo, V.D., 2013. Promptness and Academic Performance. Università di Milano Bicocca. Mimeo.
O'Donoghue, T., Rabin, M., 1999a. Doing it now or later. Am. Econ. Rev. 89 (1), 103–124.
O'Donoghue, T., Rabin, M., 1999b. Incentives for procrastinators. Q. J. Econ. 114, 769–816.
Phelps, E.S., Pollak, R., 1968. On second-best national saving and game-equilibrium growth. Rev. Econ. Stud. 35, 185–199.
Sáez-Martí, M., Sjögren, A., 2008. Deadlines and distractions. J. Econ. Theory 143, 153–176.
Schilbach, F., 2019. Alcohol and self-control: a field experiment in India. Am. Econ. Rev. 109 (4), 1290–1322.
Schraw, G., Wadkins, T., Olafson, L., 2007. Doing the things we do: a grounded theory of academic procrastination. J. Educ. Psychol. 99 (1), 12–25.
Schwartz, J., Mochon, D., Wyper, L., Maroba, J., Patel, D., Ariely, D., 2014. Healthier by precommitment. Psychol. Sci. 25 (2), 538–546.
Schwartz, J., Riis, J., Elbel, B., Ariely, D., 2012. Inviting consumers to downsize fast-food portions significantly reduces calorie consumption. Health Aff. 31 (2), 399–407.
Solomon, L., Rothblum, E., 1984. Academic procrastination: frequency and cognitive-behavioral correlates. J. Couns. Psychol. 31 (4), 503–509.
Steel, P., 2007. The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. Psychol. Bull. 133 (1), 65–94.
Taubinsky, D., 2014. From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice. Working Paper.
Taubinsky, D., Rees-Jones, A., 2018. Attention variation and welfare: theory and evidence from a tax salience experiment. Rev. Econ. Stud. 85 (4), 2462–2496.
Thaler, R.H., Benartzi, S., 2004. Save more tomorrow: using behavioral economics to increase employee saving. J. Polit. Econ. 112 (1), S164–S187.
Toussaert, S., 2018. Eliciting temptation and self-control through menu choices: a lab experiment. Econometrica 86 (3), 859–889.
Trope, Y., Fishbach, A., 2000. Counteractive self-control in overcoming temptation. J. Pers. Soc. Psychol. 79 (4), 493–506.