Working Over Time: Dynamic Inconsistency in Real Effort Tasks *

Ned Augenblick †
UC Berkeley, Haas School of Business

Muriel Niederle‡
Stanford University and NBER

Charles Sprenger§
Stanford University

First Draft: July 15, 2012
This Version: May 13, 2014

Abstract

Experimental tests of dynamically inconsistent time preferences have largely relied on choices over time-dated monetary rewards. Several recent studies have failed to find the standard patterns of present bias. However, such monetary studies contain often-discussed confounds. In this paper, we sidestep these confounds and investigate choices over consumption (real effort) in a longitudinal experiment. We pair this effort study with a companion monetary discounting study. We confirm very limited time inconsistency in monetary choices. However, subjects show considerably more present bias in effort. Furthermore, present bias in the allocation of work has predictive power for demand of a meaningfully binding commitment device. Therefore our findings validate a key implication of models of dynamic inconsistency, with corresponding policy implications.

JEL classification: C91, D12, D81

Keywords: Time Discounting, Demand for Commitment, Real Effort, Convex Time Budget

---

*We are grateful for many helpful discussions including those of Steffen Andersen, James Andreoni, Colin Camerer, Yoram Halevy, David Laibson, Matthew Rabin, and Georg Weizsäcker. We thank Wei Wu for helpful research assistance and technological expertise.

†University of California Berkeley, Haas School of Business, University of California, Berkeley, 545 Student Services Building, 1900, Berkeley, CA, 94720-1900. ned@haas.berkeley.edu

‡Stanford University, Department of Economics, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305; niederle@stanford.edu, www.stanford.edu/∼niederle

§Stanford University, Department of Economics, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305; cspreng@stanford.edu.

1 Introduction

Models of dynamically inconsistent time preferences (Strotz, 1956; Laibson, 1997; O'Donoghue and Rabin, 1999, 2001) are a pillar of modern behavioral economics, having added generally to economists' understanding of the tensions involved in consumption-savings choices, task performance, temptation, and self-control beyond the standard model of exponential discounting (Samuelson, 1937). Given the position of present-biased preferences in the behavioral literature, there is clear importance in testing the model's central falsifiable hypothesis of diminishing impatience through time. Further, testing auxiliary predictions such as sophisticated individuals' potential to restrict future activities through commitment devices can distinguish between competing accounts for behavior and deliver critical prescriptions to policy makers.[1] In this paper we present a test of dynamic inconsistency in consumption and investigate the demand for a meaningfully binding commitment device.

To date, a notably large body of laboratory research has focused on identifying the shape of time preferences (for a comprehensive review to the early 2000s, see Frederick, Loewenstein and O'Donoghue, 2002). The core of this experimental literature has identified preferences from time-dated monetary payments. Several confounds exist for identifying the shape of time preferences from such monetary choices. Issues of payment reliability and risk preference suggest that subject responses may be closely linked to their assessment of the experimenter's reliability rather than solely their time preferences.[2] Furthermore, monetary payments may not be suitable to identify parameters of models defined over time-dated consumption. Arbitrage

[1]Sophistication is taken to mean the decision-maker's recognition (perhaps partial recognition) of his predilection to exhibit diminishing impatience through time. Section 3 outlines the model which follows the framework of O'Donoghue and Rabin (2001).

[2]This point was originally raised by Thaler (1981) who, when considering the possibility of using incentivized monetary payments in intertemporal choice experiments noted 'Real money experiments would be interesting but seem to present enormous tactical problems. (Would subjects believe they would get paid in five years?)'. Recent work validates this suspicion. Andreoni and Sprenger (2012a), Gine, Goldberg, Silverman and Yang (2010), and Andersen, Harrison, Lau and Rutstrom (2012) all document that when closely controlling transactions costs and payment reliability, dynamic inconsistency in choices over monetary payments is virtually eliminated on aggregate. Further, when payment risk is added in an experimentally controlled way, non-expected utility risk preferences deliver behavior observationally equivalent to present bias as described above (Andreoni and Sprenger, 2012b).

arguments imply that choices over monetary payments should only reveal subjects' borrowing and lending opportunities (Cubitt and Read, 2007).[3] Chabris, Laibson and Schuldt (2008) describe the difficulty in mapping experimental choices over money to corresponding model parameters, casting skepticism over monetary experiments in general.

In this paper we attempt to move out of the domain of monetary choice and into the domain of consumption. Our design delivers precise point estimates on dynamic inconsistency based upon intertemporal allocations of effort and provides an opportunity to link parameter measures with demand for commitment. Delivering such a connection and contrasting present bias measured over money and over consumption are key contributions of our study.

There are few other experimental tests of dynamic inconsistency in consumption. Leading examples document dynamic inconsistency in brief, generally a few minutes, intertemporal choices over irritating noises and squirts of juice and soda (Solnick, Kannenberg, Eckerman and Waller, 1980; McClure, Laibson, Loewenstein and Cohen, 2007; Brown, Chua and Camerer, 2009). On a larger time scale, perhaps closer to everyday decision-making, there are two key contributions. Read and van Leeuwen (1998) identify dynamic inconsistency between choices over snack foods made one week apart. Ariely and Wertenbroch (2002) document demand for deadlines for classroom and work assignments, a potential sign of commitment demand for dynamically inconsistent individuals. Though suggestive, neither exercise allows for precise

[3]In a monetary discounting experiment, subjects often make binary choices between a smaller sooner payment, \$X, and a larger later payment, \$Y. The ratio, $\frac{Y}{X}$, defines a lab-offered gross interest rate. An individual who can borrow at a lower rate than the lab-offered rate should take the larger later payment, finance any sooner consumption externally, and repay their debts with the later larger payment they chose. An individual who can save at a higher rate than the lab-offered rate should take the smaller sooner payment, pay for any sooner consumption and place the remainder in their savings vehicle. These two strategies deliver a budget constraint that dominates the lab-offered budget constraint. Hence, monetary discounting experiments should reveal only external borrowing and lending opportunities. And, unless such opportunities change over time, one should reveal no present bias. The logic extends to the convex decisions of Andreoni and Sprenger (2012a). Subjects should allocate only at corner solutions and such solutions should maximize net present value at external interest rates. This point has been thoughtfully taken into account in some studies. For example, Harrison, Lau and Williams (2002) explicitly account for potential arbitrage in their calculations of individual discount rates by measuring individual borrowing and saving rates and incorporating these values in estimation. Cubitt and Read (2007) provide excellent recent discussion of the arbitrage arguments and other issues for choices over monetary payments. One counterpoint is provided by Coller and Williams (1999), who present experimental subjects with a fully articulated arbitrage argument and external interest rate information and document only a small treatment effect.

identification of discounting parameters, nor delivers the critical linkage between present bias and commitment demand. With the exception of Ashraf, Karlan and Yin (2006) and Kaur, Kremer and Mullainathan (2010) virtually no research attempts to make such links. Ashraf et al. (2006) employ monetary discounting measures and link them to take-up of a savings commitment product. Kaur et al. (2010) use disproportionate effort response on paydays to make inference on dynamic inconsistency and link this behavior to demand for a dominated daily wage contract. There are several major differences between our research and this prior work, which are discussed in detail in Section 4.4. Most importantly is the measurement of dynamic inconsistency. As opposed to problematic monetary measures or measuring only potential correlates of present bias, our effort allocations yield precise parametric measures linked directly to the theory of present bias.

102 UC Berkeley students participated in a seven week longitudinal experiment. Subjects allocated units of effort (i.e., negative leisure consumption) over two work dates. The tasks over which subjects made choices were transcription of meaningless Greek texts and completion of partial Tetris games. Allocations were made at two points in time: an initial allocation made in advance of the first work date and a subsequent allocation made on the first work date. We then randomly selected either an initial allocation or a subsequent allocation and required subjects to complete the allocated tasks. This incentivized all allocation decisions. Differences between initial and subsequent allocations allow for precise measurement of of dynamic inconsistency. A first block of the experiment, three weeks in length, was dedicated to this measurement effort.

In a second block of the experiment, also three weeks in length, the design was augmented to elicit demand for a commitment device. The commitment device of the second block allowed subjects to probabilistically favor their initial allocations over their subsequent allocations in the random selection process. Hence, choosing commitment reveals a subject's preference for implementing the allocations made in advance of the first work date. We investigate demand for our offered commitment device and correlate identified dynamic inconsistency with commitment demand.

The repeated interaction of our seven-week study allows us to complement measures of effort discounting with measures of monetary discounting taken from Andreoni and Sprenger (2012a) Convex Time Budget (CTB) choices over cash payments received in the laboratory. In these choices, subjects allocated money over two dates. Variation in whether the first payment date is the present delivers identification of monetary present bias. Hence, we can compare dynamic inconsistency measured over work and money at both the aggregate and individual level within subjects. A between-subjects replication exercise was also conducted to provide corroboration of the within-subject conclusions.

We document three primary findings. First, in the domain of money we find virtually no evidence of present bias. Monetary discount rates involving present dates are effectively indistinguishable from those involving only future dates. Further, subjects appear to treat money received at different times as perfect substitutes, suggesting they treat money as fungible. Second, in the domain of effort we find significant evidence of present bias. Subjects allocate roughly nine percent more work to the first work date when the allocation of tasks is made in advance compared to when it is made on the first work date itself. Corresponding parameter estimates corroborate these non-parametric results. Discount rates measured in advance of the first work date are around zero percent per week while discount rates measured on the first work date are around eleven percent per week. These primary study results are closely reproduced in our between-subjects replication exercise. Third, 59 percent of subjects demand commitment at price $0, preferring a higher likelihood of implementing their initial pre-work date allocations. We show that the choice of commitment is binding and meaningful in the sense that initial preferred allocations differ significantly from subsequent allocations for committing subjects. Importantly, we show that measured present bias is predictive of this commitment choice. A corresponding investigation on the extent of sophistication and commitment demand indicates that subjects potentially forecast much of their present bias. This link delivers key validation and support for our experimental measures and well-known theoretical models of present bias.

We draw two conclusions from our results. First, our results show evidence of present bias in the domain of consumption with a design that eliminates a variety of potential confounds and provides precise parameter estimation. Second, our subjects are at least partially sophisticated of their predilection to be dynamically inconsistent such that they demand binding commitment devices.

The paper proceeds as follows: Section 2 provides details for our longitudinal experimental design. Section 3 describes identification of intertemporal parameters based on experimental choices over both effort and money. Section 4 presents results. Section 5 is a discussion and section 6 concludes.

2 Design

To examine dynamic inconsistency in real effort, we introduce a longitudinal experimental design conducted over seven weeks. Subjects are asked to initially allocate tasks, subsequently allocate tasks again, and complete tasks over two work dates. Initial allocations made in advance of the first work date are contrasted with subsequent allocations made on the first work date to identify dynamic inconsistency.

If all elements of the experiment are completed satisfactorily, subjects receive a completion bonus of $100 in the seventh week of the study. Otherwise they receive only $10 in the seventh week. The objective of the completion bonus is to fix the monetary dimension of subjects' effort choices and to ensure a sizable penalty for attrition. Subjects are always paid the same amount for their work, the question of interest is *when* they prefer to complete it.

We present the design in five subsections. First, we describe the Jobs to be completed. Second, we present a timeline of the experiment and the decision environment in which allocations were made. The third subsection describes the elicitation of commitment demand. The fourth subsection addresses design details including recruitment, selection, and attrition. The fifth subsection presents the complementary monetary discounting study. In addition to this primary within-subjects study, we also conducted a between-subjects replication exercise along very similar lines. The between-subjects design is discussed primarily in section 4.5 and note is made of any design differences.

2.1 Jobs

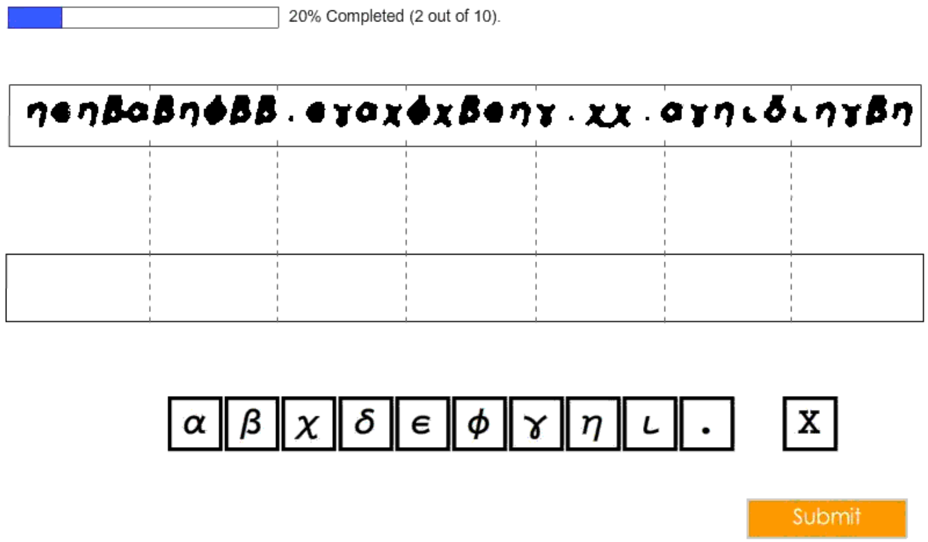
The experiment focuses on intertemporal allocations of effort for two types of job. In Job 1, subjects transcribe a meaningless Greek text through a computer interface. Panel A of Figure 1 demonstrates the paradigm. Greek letters appear in random order, slightly blurry, in subjects' transcription box. By pointing and clicking on the corresponding keyboard below the transcription box, subjects must reproduce the observed series of Greek letters. One task is the completion of one row of Greek text with 80 percent accuracy.[4] In the first week, subjects completed a task from Job 1 in an average of 54 seconds. By the final week, the average was 46 seconds.

In Job 2, subjects are asked to complete four rows of a modified Tetris game, see Panel B of Figure 1. Blocks of random shapes appear at the top of the Tetris box and fall at a fixed relatively slow speed. Arranging the shapes to complete a horizontal line of the Tetris box is the game's objective. Once a row is complete, it disappears and the shapes above fall into place. One task is the completion of four rows of Tetris. If the Tetris box fills to the top with shapes before the four rows are complete, the subject begins again with credit for the rows already completed. In the first week, subjects completed a task from Job 2 in an average of 55 seconds. By the final week, the average was 46 seconds. In contrast to a standard Tetris game, one cannot accelerate the speed of the falling shapes, and one does not pass through 'levels' of progressive difficulty. Hence, our implementation of Tetris should not be thought of as being as enjoyable as the real thing.

[4] Our measure of accuracy is the Levenshtein Distance. The Levenshtein Distance is commonly used in computer science to measure the distance between two strings and is defined as the minimum number of edits needed to transform one string into the other. Allowable edits are insertion, deletion or change of a single character. As the strings of Greek characters used in the transcription task are 35 characters long our 80 percent accuracy measure is equivalent to 7 edits or less or a Levenshtein Distance ≤ 7.

Figure 1: Experimental Jobs

Panel A: Job 1- Greek Transcription



Panel B: Job 2- Partial Tetris Games

2.2 Experimental Timeline

The seven weeks of the experiment are divided into two blocks. Weeks 1, 2, and 3 serve as the first block. Weeks 4, 5, and 6 serve as the second block. Week 7 occurs in the laboratory and is only used to pay subjects. Subjects always participate on the same day of the week throughout the experiment. That is, subjects entering the lab on a Monday allocate tasks to be completed on two future Monday work dates. Therefore, allocations are made over work dates that are always exactly seven days apart.

Weeks 1 and 4 occur in the laboratory and subjects are reminded of their study time the night before. Weeks 2, 3, 5, and 6 are completed online. For Weeks 2, 3, 5, and 6, subjects are sent an email reminder at 8pm the night before with a (subject-unique) website address. Subjects are required to log in to this website between 8am and midnight of the day in question and complete their work by 2am the following morning.

At each point of contact, subjects are first given instructions about the decisions to be made and work to be completed that day, reminded of the timeline of the experiment, given demonstrations of any unfamiliar actions, and then asked to complete the necessary actions.

The second block of the experiment, Weeks 4, 5, and 6, mimics the first block of Weeks 1, 2, and 3, with one exception. In Week 4, subjects are offered a probabilistic commitment device, which is described in detail in subsection 2.4. Hence, we primarily describe Weeks 1, 2 and 3 and note any design changes for Weeks 4, 5 and 6. To summarize our longitudinal effort experiment, Table 1 contains the major events in each week which are described in detail below.

Table 1: Summary of Longitudinal Experiment

	10 Effort Allocations	Minimum Work	Allocation-That-Counts Chosen	Complete Work	Commitment Choice	Receive Payment
Week 1 (In Lab):	x	x				
Week 2 (Online):	x	x	x	x		
Week 3 (Online):		x		x		
Week 4 (In Lab):	x	x			x	
Week 5 (Online):	x	x	x	x		
Week 6 (Online):		x		x		
Week 7 (In Lab):						x

2.3 Effort Allocations

In Week 1, subjects allocate tasks between Weeks 2 and 3. In Week 2, subjects also allocate tasks between Weeks 2 and 3. Subjects were not reminded of their initial Week 1 allocations in

Week 2. Note that in Week 1 subjects are making decisions involving two future work dates, whereas in Week 2, subjects are making decisions involving a present and a future work date. Before making decisions in Week 1, subjects are told of the Week 2 decisions and are aware that exactly one of all Week 1 and Week 2 allocation decisions will be implemented.

2.3.1 Allocation Environment

Allocations are made in a convex decision environment. Using slider bars, subjects allocate tasks to two dates, one earlier and one later, under different interest rates.[5] Figure 2 provides a sample allocation screen. To motivate the intertemporal tradeoffs faced by subjects, decisions are described as having different 'task rates' such that every task allocated to the later date reduces the number of tasks allocated to the sooner date by a stated number. For example, a task rate of 1:0.5 implies that each task allocated to Week 3 reduces by 0.5 the number allocated to Week 2.[6]

For each task and for each date where allocations were made, subjects faced five task rates. These task rates take the values, $R \in \{0.5, 0.75, 1, 1.25, 1.5\}$. The subjects' decision can be formulated as allocating tasks e over times t and $t+k$, e_t and e_{t+k}, subject to the present-value budget constraint,

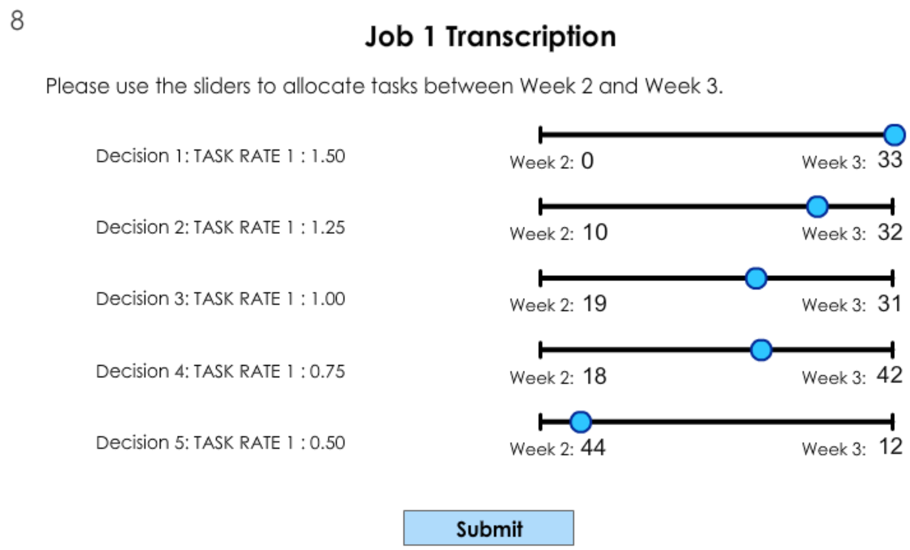
$$e_t + R \cdot e_{t+k} = m. \tag{1}$$

The number of tasks that subjects could allocate to the sooner date was capped at fifty such that $m = 50$ in each decision in the experiment.

[5]The slider was initially absent from each slider bar and appeared in the middle of the bar once a subject clicked on the allocation. Every slider bar was thus clicked on before submission, avoiding purely passive response.

[6]We thank an anonymous referee for noting a small error in our instructions which inverted the task rates when first introducing them. Though this appears not to have affected response as allocations move appropriately with task rates, we do correct this error in our replication exercise and document very similar behavior. See section 4.5 for detail.

Figure 2: Convex Allocation Environment



2.3.2 Minimum Work

In each week, subjects are required to complete 10 tasks of each Job prior to making allocation decisions or completing allocated tasks. The objective of these required tasks, which we call "minimum work," is three-fold. First, minimum work requires a few minutes of participation at each date, forcing subjects to incur the transaction costs of logging on to the experimental website at each time.[7] Second, minimum work, especially in Week 1, provides experience for subjects such that they have a sense of how effortful the tasks are when making their allocation decisions. Third, we require minimum work in all weeks before all decisions, and subjects are informed that they will have to complete minimum work at all dates. This ensures that subjects have experienced and can forecast having experienced the same amount of minimum work when making their allocation decisions at all points in time.

[7]A similar technique is used in monetary discounting studies where minimum payments are employed to eliminate subjects loading allocations to certain dates to avoid transaction costs of receiving multiple payments or cashing multiple checks (Andreoni and Sprenger, 2012a).

2.3.3 The Allocation-That-Counts

Each subject makes 20 decisions allocating work to Weeks 2 and 3: five decisions are made for each Job in Week 1 and five for each Job in Week 2. After the Week 2 decisions, one of these 20 allocations is chosen at random as the 'allocation-that-counts' and subjects have to complete the allocated number of tasks on the two work dates to ensure successful completion of the experiment (and hence payment of $100 instead of only $10 in Week 7).

The randomization device probabilistically favors the Week 2 allocations over the Week 1 allocations. In particular, subjects are told (from the beginning) that their Week 1 allocations will be chosen with probability 0.1, while their Week 2 allocations will be chosen with probability 0.9. Within each week's allocations, every choice is equally likely to be the allocation-that-counts.[8] This randomization process ensures incentive compatibility for all decisions. This design choice was made for two reasons. First, it increases the chance that subjects experienced their own potentially present-biased behavior. Second, it provides symmetry to the decisions in Block 2 that elicit demand for commitment.

2.4 Commitment Demand

In the second block of the experiment, Weeks 4, 5, and 6, subjects are offered a probabilistic commitment device. In Week 4, subjects are given the opportunity to choose which allocations will be probabilistically favored. In particular, they can choose whether the allocation-that-counts comes from Week 4 with probability 0.1 (and Week 5 with probability 0.9), favoring flexibility, or from Week 4 with probability 0.9, favoring commitment. This form of commitment device was chosen because of its potential to be meaningfully binding. Subjects who choose to commit and who differ in their allocation choices through time can find themselves constrained by commitment with high probability.

In order to operationalize our elicitation of commitment demand, subjects are asked to make 15 multiple price list decisions between two options. In the first option, the allocation-

[8]For the description of the randomization process given to subjects please see instructions in Appendix D.

that-counts will come from Week 4 with probability 0.1. In the second option, the allocation-that-counts will come from Week 4 with probability 0.9. In order to determine the strength of preference, an additional payment of between \$0 and \$10 is added to one of the options for each decision.[9] Figure 3 provides the implemented price list. One of the 15 commitment decisions is chosen for implementation, ensuring incentive compatibility. Subjects are told that the implementation of the randomization for the commitment decisions will occur once they submit their Week 5 allocation decisions. Given this randomization procedure, an individual choosing commitment in all 15 decisions will complete a Week 4 allocation with probability 0.9. Each row at which a subject chooses flexibility reduces this probability by 5.3 percent.[10] Hence a subject choosing to commit at price zero (the eighth row) and lower will complete an initial allocation with probability 0.53. Naturally, if subjects treat each commitment decision in isolation, the incentives are more stark as each decision moves the probability of facing an initial allocation from 0.1 to 0.9.[11] This isolation is encouraged in the design as subjects are told to treat each commitment decision as if it was the one going to be implemented (See Appendix D.4 for detail).

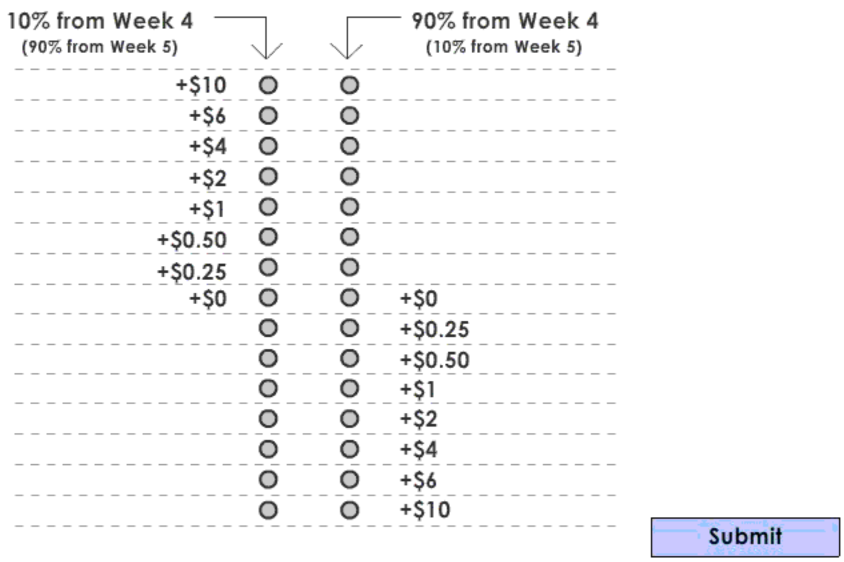
Our commitment demand decisions, and the second block of the experiment, serve three purposes. First, they allow us to assess the demand for commitment and flexibility. Second, a key objective of our study is to explore the theoretical link, under the assumption of sophistication, between present bias and commitment demand. Are subjects who are present-biased more likely to demand commitment? Third, a correlation between time inconsistency and commitment validates the interpretation of present bias over other explanations for time inconsistent effort choices. For example, a subject who has a surprise exam in Week 2 may be observationally indistinguishable in her Week 2 effort choices from a present-biased subject. However, a subject prone to surprise exams should favor flexibility to accommodate her noisy schedule.

[9] We chose not to have the listed prices ever take negative values (as in a cost) to avoid subjects viewing paying for commitment as a loss.

[10] Each row changes the probability of implementing an initial allocation by $(1/15 * (0.9 - 0.1)) = 0.053$.

[11] In assessing the value of commitment we make this assumption, ignoring the second stage randomization inherent to the commitment demand elicitation.

Figure 3: Commitment Demand Elicitation



On the other hand, a sophisticated present-biased subject may demand commitment to restrict her future self.

2.5 Design Details

102 UC Berkeley student subjects were initially recruited into the experiment across 4 experimental sessions on February 8th, 9th and 10th, 2012 and were told in advance of the seven week longitudinal design and the $100 completion bonus.[12] Subjects did not receive an independent show up fee. 90 subjects completed all aspects of the working over time experiment and received the $100 completion bonus. The 12 subjects who selected out of the experiment do not appear different on either initial allocations, comprehension or a small series of demographic data collected at the end of the first day of the experiment.[13] One more subject completed initial allocations in Week 1, but due to computer error did not have their choices recorded.

[12]Student subjects were recruited from the subject pool of the UC Berkeley Experimental Laboratory, Xlab. Having subjects informed of the seven week design and payment is a potentially important avenue of selection. Our subjects were willing to put forth effort and wait seven weeks to receive $100. Though we have no formal test, this suggests that our subjects may be a relatively patient selection.

[13]3 of those 12 subjects dropped after the first week while the remaining 9 dropped after the second week. Including data for these 9 subjects where available does not qualitatively alter the analysis or conclusions.

This leaves us with 89 subjects.

One critical aspect of behavior limits our ability to make inference for time preferences based on experimental responses. In particular, if subjects have no variation in allocations in response to changes in R in some weeks, then attempting to point identify both discounting and cost function parameters is difficult, yielding imprecise and unstable estimates. Similar to multiple price list experiments, if a subject always chooses a specific option, only one-sided bounds on parameters can be obtained. Here, the problem is compounded by our efforts to identify both discounting and cost function parameters. In our sample, nine subjects have this issue for one or more weeks of the study. For the analysis, we focus on the primary sample of 80 subjects who completed all aspects of the experiment with positive variation in their responses in each week. In Appendix Table A2, we re-conduct the aggregate analysis including these nine subjects and obtain very similar findings.

2.6 Monetary Discounting

Subjects were present in the laboratory in the first, fourth, and seventh week of the experiment. This repeated interaction facilitates a monetary discounting study that complements our main avenue of analysis. In Weeks 1 and 4 of our experimental design, once subjects complete their allocation of tasks, they are invited to respond to additional questions allocating monetary payments to Weeks 1, 4, and 7. In Week 1, we implement three Andreoni and Sprenger (2012a) Convex Time Budget (CTB) choice sets, allocating payments across: 1) Week 1 vs. Week 4; 2) Week 4 vs. Week 7 (Prospective); and 3) Week 1 vs. Week 7. Individuals are asked to allocate monetary payments c across the two dates t and $t+k$, c_t and c_{t+k}, subject to the intertemporal budget constraint,

$$P \cdot c_t + c_{t+k} = m. \tag{2}$$

The experimental budget is fixed at $m = \$20$ and five interest rates are implemented in each choice set, summarized by $P \in \{0.99, 1, 1.11, 1.25, 1.43\}$. These values were chosen for comparison with prior work (Andreoni and Sprenger, 2012a).[14] In Week 4, we ask subjects to allocate in a CTB choice set over Week 4 and Week 7 under the same five values of P. We refer to these choices made in Week 4 as Week 4 vs. Week 7 and those made in Week 1 over these two dates as Week 4 vs. Week 7 (Prospective). Hence, subjects complete a total of four CTB choice sets.

The CTBs implemented in Weeks 1 and 4 are paid separately and independently from the rest of the experiment with one choice from Week 1 and one choice from Week 4 chosen to be implemented. Subjects are paid according to their choices. Subjects are not told of the Week 4 choices in Week 1. As in Andreoni and Sprenger (2012a), we have minimum payments of \$5 at each payment date to ensure equal transaction costs in each week, such as waiting to get paid. Appendix D provides the full experimental instructions.

The implemented monetary discounting experiments have two nuances relative to Andreoni and Sprenger (2012a). First, Andreoni and Sprenger (2012a) implement CTBs with payment by check. Our design implements payment by cash with potentially lower transaction costs. Second, Andreoni and Sprenger (2012a) implement CTBs with present payment received only by 5:00 p.m. in a subject's residence mailbox. If these payments are not construed as being received in the present, one would expect no present bias. Here, we provide payment immediately in the laboratory.

In both Weeks 1 and 4, the monetary allocations are implemented after the more central effort choices. The monetary choices were not announced in advance and subjects could choose not to participate; five did so in either Weeks 1 or 4. In our analysis of monetary discounting, we focus on the 75 subjects from the primary sample with complete monetary choice data.

3 Theoretical Structure and Identification

In the intertemporal allocation of effort and money, discounting and additional parameters can be identified at either the aggregate or individual level under various structural assumptions. In the following three subsections we describe our theoretical environment, explore the demand

[14]Additionally, $P = 0.99$ allows us to investigate the potential extent of negative discounting.

for commitment, demonstrate which experimental variation provides identification of specific parameters of interest, and lay out methodology for estimation. A fourth subsection presents estimation details for monetary discounting.

3.1 Effort Discounting

3.1.1 Allocation Timing

In the working over time experiment, subjects allocate effort to an earlier date, e_t, and a later date, e_{t+k}, subject to the intertemporal budget constraint described in (1). Subjects make allocations at two points in time, one at time $s < t$, and one at time t. The allocation-that-counts is randomly implemented from time s with probability p and from time t with probability $1-p$.[15] Let $e_{t,s}$ be the allocation of effort to time t chosen at time s. Let $e^{s*}_{t,t}$ be the allocation of effort to time t forecasted to be chosen at time t from the perspective of time s. That is, $e^{s*}_{t,t}$ captures what an individual at time s believes they will optimally choose at time t.

3.1.2 Preferences

To develop our theory, we assume an instantaneous cost function, $c(e)$, for effort, e, that is time separable, stationary, and of an expected utility form with respect to the probability that an allocation is implemented. To aid our development and foreshadow our empirical implementation we also make a functional form assumption for the shape of $c(\cdot)$. We assume

$$c(e) = (e + \omega)^{\gamma},$$

where $\gamma > 1$ represents the stationary parameter on the convex instantaneous cost of effort function. The additive term ω in the cost function could be interpreted as a Stone-Geary minimum or as some background level of required work. Such parameters are used in monetary discounting studies (Andersen, Harrison, Lau and Rutstrom, 2008; Andreoni and Sprenger,

[15]We abstract from the fact that subjects make multiple allocations. Given the assumed separability over time and in probabilities, this abstraction is innocuous.

2012a), and are either taken from some external data source on background consumption or estimated from experimental choices. For simplicity, we interpret ω as the required minimum work of the experiment and set $\omega = 10$ for our effort analysis.[16]

We assume discounting follows the quasi-hyperbolic partially sophisticated form proposed by O'Donoghue and Rabin (2001). For two periods, t and $t+k$, discounting, $D(t, t+k)$, is captured by

$$D(t, t+k) = \begin{cases} \beta\delta^k & \text{if } k > 0 \\ 1 & \text{if } k = 0. \end{cases}$$

The parameter β captures the degree of present bias while the parameter δ captures long run discounting. $\beta = 1$ nests the standard model of exponential discounting. From period $s < t$, the discounted costs of effort at times t and $t+k$ can be written as

$$\beta\delta^{t-s}(e_{t,s} + \omega)^\gamma + \beta\delta^{t+k-s}(e_{t+k,s} + \omega)^\gamma.$$

Eliminating common terms, the decision problem at time s can be written as

$$\begin{aligned} min_{e_{t,s} \cdot e_{t+k,s}} \quad & p \cdot [(e_{t,s} + \omega)^\gamma + \delta^k(e_{t+k,s} + \omega)^\gamma] + \\ & (1-p) \cdot [(e^{s*}_{t,t} + \omega)^\gamma + \delta^k(e^{s*}_{t+k,t} + \omega)^\gamma] \\ s.t. \quad & e_{t,s} + R \cdot e_{t+k,s} = m, \end{aligned}$$

which yields the intertemporal Euler equation satisfied by the optimal allocation, $(e^*_{t,s}, e^*_{t+k,s})$,

$$\left(\frac{e^*_{t,s} + \omega}{e^*_{t+k,s} + \omega}\right)^{\gamma-1} \frac{1}{\delta^k} = \frac{1}{R}.$$

Note that the forecasted allocation, $(e^{s*}_{t,t}, e^{s*}_{t+k,t})$, and the probability of implementation, p, do not feature in the intertemporal Euler due to the assumed separability. Similarly, the decision

[16] Andreoni and Sprenger (2012a) provide estimates for ω based on non-linear least squares techniques and analyze the extent to which different assumptions for ω influence remaining parameter estimates. Though utility curvature and discounting are sensitive to varying assumptions for ω, present bias, β, is largely unaffected. Andersen et al. (2008) also provide some sensitivity analysis.

problem at time t can be written

$$\begin{aligned} min_{e_{t,t}.e_{t+k,t}} \quad & p \cdot [(e^*_{t,s} + \omega)^\gamma + \beta\delta^k (e^*_{t+k,s} + \omega)^\gamma] + \\ & (1-p) \cdot [[(e_{t,t} + \omega)^\gamma + \beta\delta^k (e_{t+k,t} + \omega)^\gamma]] \\ s.t. \quad & e_{t,t} + R \cdot e_{t+k,t} = m, \end{aligned}$$

with corresponding Euler equation satisfied by the optimal allocation, $(e^*_{t,t}, e^*_{t+k,t})$,

$$(\frac{e^*_{t,t} + \omega}{e^*_{t+k,t} + \omega})^{\gamma-1} \frac{1}{\beta\delta^k} = \frac{1}{R}$$

The prior allocation, $(e^*_{t,s}, e^*_{t+k,s})$, and the probability of implementation do not feature in the intertemporal Euler. Any differences in allocations between time s and time t are delivered by the present bias term, β.[17]

Combining our Euler equations we have

$$(\frac{e^*_{t,D} + \omega}{e^*_{t+k,D} + \omega})^{\gamma-1} \frac{1}{\beta^{\mathbf{1}_{D=t}} \delta^k} = \frac{1}{R} \tag{3}$$

where $D \in \{s, t\}$ represents whether the allocation decision was made at time t or time s. Note that for $\beta < 1$, an allocation made at time t at a given R will have a lower value of $e^*_{t,D}$ than

[17]A recent discussion of non-expected utility behavior in intertemporal settings has demonstrated that apparently present-biased behavior can be delivered by deviations from expected utility (see, e.g., Halevy, 2008). Under discounted expected utility, allocations over two periods should depend on the ratio of probabilities with which the allocations are realized. In two important conditions Andreoni and Sprenger (2012b) demonstrate in the monetary domain that if sooner and later payments are paid independently with probability 0.5, behavior deviates from the common ratio counterpart of all payments being certain. Under expected utility and sensible formulations of prospect theory, the deviations cannot be rationalized. An intuition for the effect is that the independent payment probabilities give subjects the opportunity to hedge through time. Cheung (Forthcoming) and Miao and Zhong (2012) demonstrate the importance of this intuition, as they show in the Andreoni and Sprenger (2012b) setup that when one makes the two 0.5 realization probabilities perfectly correlated behavior is closer to the expected utility benchmark. In our environment, the implementation probability applies equally to both the sooner and later work date, creating perfect correlation through time. Hence, the effects of Andreoni and Sprenger (2012b) are unlikely to be present. Additionally, because the same implementation probability applies to both work dates any non-linear treatment of p or $1-p$ must be applied equally, and so drop out of marginal conditions in exactly the same way that undistorted probabilities do. Further potential concerns with respect to the asymmetry of p and $1-p$ in the design are addressed in our replication exercise where initial and subsequent allocations are implemented with equal probability. See section 4.5 for detail.

an allocation made at time s. A present-biased individual allocates less work to time t at time t than they did at time s.

Naturally, the prediction that dynamically inconsistent behavior depends only on β relies on the assumption of a stationary cost function. Changes in the cost function through time could easily lead to differences in allocations between time s and t. Such changing costs could be delivered by a variety of sources. For example, there could be permanent shocks to the cost function, perhaps due to a misforecasting of task difficulty. There could also be temporary shocks due to some random events that impose time constraints or leave subjects more tired and exhausted than they normally are. In section 5 we address these concerns directly and provide evidence that such possibilities are unlikely to drive observed behavior.

3.1.3 Partial Sophistication

We allow for the fact that individuals may be partially sophisticated with respect to their own present bias. The nature of sophistication follows that of O'Donoghue and Rabin (2001), where $\widehat{\beta}$ captures the belief an individual has on his future present bias: $\widehat{\beta} = \beta$ represents full sophistication, $\widehat{\beta} = 1$ represents full naivete, and $\widehat{\beta} \in (\beta, 1)$ represents partial sophistication. This means that allocations at time t, forecasted at time $s < t$ are

$$
\begin{aligned}
(e^{s*}_{t,t}, e^{s*}_{t+k,t}) = \quad & argmin\ p \cdot [(e^*_{t,s} + \omega)^\gamma + \widehat{\beta}\delta^k(e^*_{t+k,s} + \omega)^\gamma] + \\
& (1-p) \cdot [(e^s_{t,t} + \omega)^\gamma + \widehat{\beta}\delta^k(e^s_{t+k,t} + \omega)^\gamma] \\
s.t. \quad & e^s_{t,t} + R \cdot e^s_{t+k,t} = m.
\end{aligned}
$$

If $\widehat{\beta} \in (\beta, 1]$ an individual's forecasted allocation, $(e^{s*}_{t,t}, e^{s*}_{t+k,t})$ will not accord with their actual subsequent allocation, $(e^*_{t,t}, e^*_{t+k,t})$.

Note the sophistication parameter, $\widehat{\beta}$ is absent from the Euler formulations above. This is by construction both in the theory and the experimental design. An individual at time s may forecast a level of present bias at time t but is incapable of controlling behavior at that point in time. More importantly, this forecasted present bias at time t does not influence his behavior at time s. The only actions available to the time t self is to complete the time s allocation with probability p, complete the time t allocation with probability $1-p$, or opt out of the experiment, foregoing \$90. Given the high penalty, an individual at time s can appropriately forecast the third action will not be taken. The individual is aware that he cannot control the second action. Hence, he optimizes according to his time s preference as above with $\widehat{\beta}$ absent from the formulation. The parameter $\widehat{\beta}$ will be important for our analysis of commitment in which an individual at time s may indeed control time t behavior.

3.2 Commitment

In the second block of the experiment subjects are offered a probabilistic commitment device. The commitment device favors the initial allocations made at time s over the subsequent allocations made at time t by changing the time s implementation probability from p to $1-p$ (i.e. from 0.1 to 0.9).

Recall that intertemporal Euler equations and allocations are independent of implementation probabilities. Hence, the value of commitment can be arrived at by comparing discounted costs. An individual prefers to commit if the discounted costs of the chosen allocation at time s are smaller than the discounted costs of the forecasted allocation for time t at time s.[18] The

[18] The inequality between discounted costs

$$(1-p)\cdot[\beta\delta^{t-s}(e^*_{t,s}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^*_{t+k,s}+\omega)^\gamma]+p\cdot[\beta\delta^{t-s}(e^{s*}_{t,t}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^{s*}_{t+k,t}+\omega)^\gamma]<$$
$$p\cdot[\beta\delta^{t-s}(e^*_{t,s}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^*_{t+k,s}+\omega)^\gamma]+(1-p)\cdot[\beta\delta^{t-s}(e^{s*}_{t,t}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^{s*}_{t+k,t}+\omega)^\gamma],$$

reduces to the inequality,

$$(e^*_{t,s}+\omega)^\gamma+\delta^k c(e^*_{t+k,s}+\omega)^\gamma<(e^{s*}_{t,t}+\omega)^\gamma+\delta^k c(e^{s*}_{t+k,t}+\omega)^\gamma,$$

provided $p<0.5$ as in the experiment. Subtracting the discounted costs one arrives at the value of commitment,

$$V=\{p\cdot[\beta\delta^{t-s}(e^*_{t,s}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^*_{t+k,s}+\omega)^\gamma]+(1-p)\cdot[\beta\delta^{t-s}(e^{s*}_{t,t}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^{s*}_{t+k,t}+\omega)^\gamma]\}-$$
$$\{(1-p)\cdot[\beta\delta^{t-s}(e^*_{t,s}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^*_{t+k,s}+\omega)^\gamma]+p\cdot[\beta\delta^{t-s}(e^{s*}_{t,t}+\omega)^\gamma+\beta\delta^{t+k-s}c(e^{s*}_{t+k,t}+\omega)^\gamma]\}.$$

value of commitment is given as

$$V = (1-2p)\cdot\beta\delta^{t-s}\cdot\{[(e^{s*}_{t,t}+\omega)^{\gamma}+\delta^{k}c(e^{s*}_{t+k,t}+\omega)^{\gamma}]-[(e^{*}_{t,s}+\omega)^{\gamma}+\delta^{k}c(e^{*}_{t+k,s}+\omega)^{\gamma}]\}.$$

Note that the value of commitment, V, depends upon both actual allocations and forecasted allocations at time s. Hence, the value of commitment depends upon the degree of sophistication. Clearly, for naive individuals with $\widehat{\beta}=1$, $(e^{s*}_{t,t}, e^{s*}_{t+k,t}) = (e^{*}_{t,s}, e^{*}_{t+k,s})$. Actual and forecasted allocations are identical and the value of commitment is zero.

For sophisticated individuals, $\widehat{\beta}\in[\beta,1)$, actual allocations and forecasted allocations at time s differ. By the definition of the minimum from the perspective of period s, $(e^{s*}_{t,t}, e^{s*}_{t+k,t})$ yields higher discounted costs than $(e^{*}_{t,s}, e^{*}_{t+k,s})$. This implies that the value of commitment should be positive provided $p<0.5$, as in the experiment. As $\widehat{\beta}$ diverges from 1, the value of commitment increases. Appendix C provides further detail and corresponding simulated values. The extent of commitment demand, when combined with parametric measures for discounting and costs, can be informative for the extent of sophistication.

Naturally, there may be intrinsic benefits to flexibility. These unmodeled benefits to flexibility could have many sources including future uncertainty in costs or task difficulty.[19]

The value of commitment, V, is measured in the same units as the discounted costs of effort. A potential shortfall of our design is that our experiment does not measure V directly but rather measures its translation into dollars. Hence, we provide potential bounds on V based upon assumptions for the transformation of V to dollars.

[19]Note that in the presence of such factors even sophisticated present-biased subjects may have low or even negative values for commitment. Hence, it is critical that our design elicits the demand for both flexibility and commitment to assess the possible presence of such factors.

3.3 Identification

From the intertemporal Euler equation, (3), identification of discounting and the cost function is straightforward. Rearranging and taking logs yields

$$log(\frac{e_{t,D}+\omega}{e_{t+k.D}+\omega}) = \frac{log(\beta)}{\gamma-1}\cdot(\mathbf{1}_{D=t}) + \frac{log(\delta)}{\gamma-1}\cdot k - (\frac{1}{\gamma-1})\cdot log(R), \qquad (4)$$

which is linear in the key experimental parameters of whether allocations are made at time t, $\mathbf{1}_{D=t}$, and the log transform, $log(R)$. In our implementation, variation in $log(R)$ delivers identification of the cost function, γ; the allocation being made in Week 1 ($D = s$) rather than Week 2 ($D = t$) delivers identification of present bias, β; and the delay length, $k = 7$ days, gives identification of the discount factor, δ.[20]

In order to estimate discounting and cost function parameters from aggregate data, we assume an additive error structure and estimate the linear regression implied by (4). The parameters of interest can be recovered from non-linear combinations of regression coefficients with standard errors calculated via the delta method.[21] One important issue to consider in the estimation of (4) is the potential presence of corner solutions. We provide estimates from two-limit tobit regressions designed to account for the possibility that the tangency condition implied by (4) does not hold with equality (Wooldridge, 2002).

Estimating (4) is easily extended to the study of individual parameters. To begin, (4) can be estimated at the individual level.[22] However, with limited numbers of individual choices it is helpful to consider alternative, more structured approaches. In particular, we allow for heterogeneous discounting across individuals, but assume all individuals have the same cost

[20] Of course, with only one delay length of seven days considered in the experiment, we have limited confidence that our estimate of δ can be extrapolated to arbitrary delay lengths.

[21] To be specific, the regression equation is, for $k = 7$,

$$log(\frac{e_t+\omega}{e_{t+k}+\omega})_i = \eta_0 k + \eta_1\cdot(\mathbf{1}_{D=t})_i + \eta_2\cdot log(R)_i + \epsilon_i,$$

and we recover the parameters of interest as $\beta = exp(\hat{\eta}_1/-\hat{\eta}_2)$ and $\gamma = 1+1/-\hat{\eta}_2$. Note that $\hat{\delta} = exp(\hat{\eta}_0/-\hat{\eta}_2)$ is recovered from the constant as only one delay length was used in the experimental design.

[22] Broadly similar conclusions are reached when estimating (4) at the individual level, however, parameter precision is greatly reduced and substantial estimate instability is uncovered in some cases.

function. Consider a vector of fixed effects $(\mathbf{1}_j)_i$ which take the value 1 if observation i was contributed by individual j. This leads to the fixed effects formulation

$$log(\frac{e_{t,D}+\omega}{e_{t+k,D}+\omega})_i = \frac{log(\overline{\delta})}{\gamma-1}\cdot k + \frac{(log(\boldsymbol{\delta}_j)-log(\overline{\delta}))}{\gamma-1}\cdot(\mathbf{1}_j)_i\cdot k + \frac{log(\overline{\beta})}{\gamma-1}\cdot(\mathbf{1}_{D=t})_i$$
$$+\frac{(log(\boldsymbol{\beta}_j)-log(\overline{\beta}))}{\gamma-1}\cdot(\mathbf{1}_{D=t})_i\cdot(\mathbf{1}_j)_i - \frac{1}{\gamma-1}\cdot log(R)_i,$$

where $\overline{\delta}$, $\overline{\beta}$ refer to sample means, and δ_j, β_j refer to individual-specific discounting parameters. With an additive error structure this is easily estimable.[23] The individual fixed effect interacted with the decision being made in the present provides identification of the individual-specific β_j. In Appendix A we conduct simulation exercises under various correlation structures for the true parameters of interest and document that the implemented estimation methods perform well both at the aggregate and individual level.

3.4 Monetary Discounting

Our methods for recovering monetary discounting parameters at both the aggregate and individual level closely follow those for effort. Following most of the literature, we abstract from standard arbitrage arguments for monetary discounting and assume laboratory administered rates are the relevant ones.[24] In particular, for monetary payments, c_t and c_{t+k}, allocated subject to the constraint (2), we assume a quasi-hyperbolic constant relative risk averse utility function,

$$U(c_{t,D}, c_{t+k,D}) = (c_t+\omega)^\alpha + \beta^{\mathbf{1}_{D=t}}\delta^k(c_{t+k}+\omega)^\alpha. \tag{5}$$

Where $D \in \{s,t\}$ refers to the same notation as before for when the allocation decision is made. The utility function is assumed to be concave, $\alpha < 1$, such that first order conditions provided

[23]We allow both β and δ to vary across individuals such that the implemented regression is a standard interaction with both level and slope effects.

[24]The assumptions that individuals narrowly bracket time-dated experimental payments, treat money effectively as consumption, and ignore extra-lab arbitrage have been standard in the literature. One prominent exception to this tradition is Harrison et al. (2002), who measure and account for extra-lab borrowing and savings opportunities.

meaningful optima. Here, the parameter ω is a background parameter that we take to be the \$5 minimum payment of the monetary experiment.[25]

Maximizing (6) subject to the intertemporal budget constraint (2) yields an intertemporal Euler equation similar to that above for effort. Taking logs and rearranging we have

$$log(\frac{c_{t,D}+\omega}{c_{t+k,D}+\omega}) = \frac{log(\beta)}{\alpha-1}\cdot(\mathbf{1}_{D=t}) + \frac{log(\delta)}{\alpha-1}\cdot k + (\frac{1}{\alpha-1})\cdot log(P), \tag{6}$$

which can again be estimated at the aggregate or individual level via two-limit Tobit. Discounting and utility function parameters can be recovered via non-linear combinations of regression coefficients as above with standard errors estimated again via the delta method.

4 Results

The results are presented in four subsections. First, we present aggregate results from the monetary discounting study and compare our observed level of limited present bias with other recent findings. Second, we move to effort related discounting and provide both non-parametric and parametric aggregate evidence of present bias. Third, we analyze individual heterogeneity in discounting for both work and money. Lastly, in a fourth subsection we present results related to commitment demand, document correlations with previously measured present bias and analyze the value of commitment.

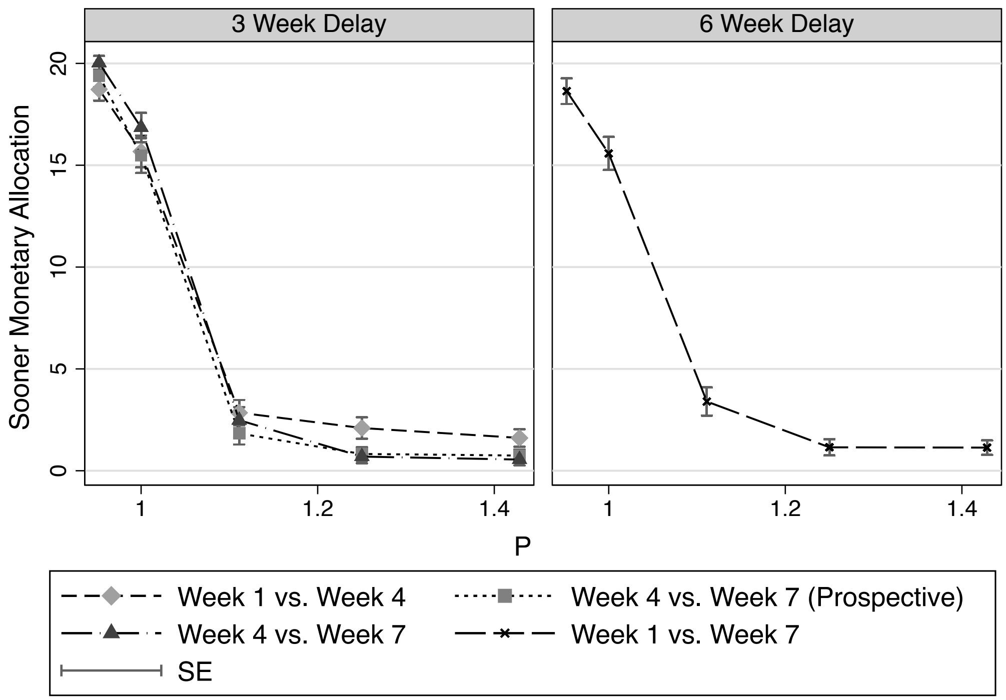
4.1 Monetary Discounting

Figure 4 presents the data from our monetary discounting experiment. The mean allocation to the sooner payment date at each value of P is reported for the 75 subjects from the primary sample for whom we have all monetary discounting data. The left panel shows three data series

[25] Andreoni and Sprenger (2012a) provide detailed discussion of the use of such background parameters and provide robustness tests with differing values of ω and differing assumptions for the functional form of utility in CTB estimates. The findings suggest that though utility function curvature estimates may be sensitive to different background parameter assumptions, discounting parameters, particularly present bias, are virtually unaffected by such choices.

for payments sets with three-week delay lengths while the right panel shows the data series for the payment sets with a six-week delay length. Standard error bars are clustered at the individual level.

Figure 4: Monetary Discounting Behavior



We highlight two features of Figure 4. First, note that as P increases, the average allocation to the sooner payment decreases, following the law of demand. Indeed, at the individual level 98% of choices are monotonically decreasing in P, and only 1 subject exhibits more than 5 non-monotonicities in demand in their monetary choices.[26] This suggests that subjects as a whole understand the implied intertemporal tradeoffs and the decision environment.

Second, Figure 4 allows for non-parametric investigation of present bias in two contexts.[27]

[26]Subjects have 16 opportunities to violate monotonicity comparing two adjacent values of P in their 20 total CTB choices. 63 of 75 subjects have no identified non-monotonocities. Andreoni and Sprenger (2012a) provide a detailed discussion of the extent of potential errors in CTB choices. In particular they note that prevalence of non-monotonicities in demand are somewhat less than the similar behavior of multiple switching in standard Multiple Price List experiments.

[27]Though the six-week delay data are used in estimation, our non-parametric tests only identify present bias

First, one can consider the static behavior, often attributed to present bias, of subjects being more patient in the future than in the present by comparing the series Week 1 vs. Week 4 and Week 4 vs. Week 7 (Prospective). In this comparison, controlling for P, subjects allocate on average \$0.54 ($s.e = 0.31$) more to the sooner payment when it is in the present, $F(1, 74) = 2.93,\ (p = 0.09)$. A second measure of present bias is to compare Week 4 vs. Week 7 (Prospective) made in Week 1 to the Week 4 vs. Week 7 choices made in Week 4. This measure is similar to the recent work of Halevy (2012). Ignoring income effects associated with having potentially received prior experimental payments, this comparison provides a secondary measure of present bias. In this comparison, controlling for P, subjects allocate on average \$0.47 ($s.e = 0.32$) more to the sooner payment when the sooner payment is in the present, $F(1, 74) = 2.08,\ (p = 0.15)$.[28]

Over monetary payments, we find limited non-parametric support for the existence of a present bias. In Table 2, columns (1) and (2) we provide corresponding parameter estimates implementing two-limit Tobit regressions of (6), with standard errors clustered at the individual level. In column (1) we use all 4 CTB choice sets. In column (2) we use only the choice sets which have three-week delays for continuity with both our non-parametric evidence and the comparisons generally made in experimental economics. Across specifications we identify daily discount factors of around 0.998. The 95% confidence interval in column (1) for the daily discount factor implies annual discount rates between 40% and 140%.[29] In column (1) of Table 2 we estimate $\beta = 0.974$ ($s.e. = 0.009$), economically close to, though statistically different from dynamic consistency, $H_0 : \beta = 1$: $\chi^2(1) = 8.77,\ (p < 0.01)$. In column (2), focusing only on three week delay data, we find $\beta = 0.988$ (0.009) and are unable to reject the null hypothesis

from choices over three-week delays. We ignore here the method of identifying present bias frequently used in psychology where short horizon choices are compared to long horizon choices.

[28] Additionally, this measure is close in spirit to our effort experiment where initial allocations are compared to subsequent allocations. To get a sense of the size of potential income effects, we can also compare the Week 1 vs. Week 4 choices made in Week 1 to the Week 4 vs. Week 7 choices made in Week 4. Controlling for P, subjects allocate on average \$0.07 ($s.e = 0.31$) more to the sooner payment in Week 1, $F(1, 74) = 0.05,\ (p = 0.82)$, suggesting negligible income effects.

[29] Admittedly, our ability to precisely identify aggregate discounting was not a focus of the experimental design and is compromised by limited variation in delay length and interest rates. In monetary discounting experiments it is not unusual to find implied annual discount rates in excess of 100%.

Table 2: Parameter Estimates

	Monetary Discounting		Effort Discounting		
	(1) All Delay Lengths	(2) Three Week Delay Lengths	(3) Job 1 Greek	(4) Job 2 Tetris	(5) Combined
Present Bias Parameter: β	0.974	0.988	0.900	0.877	0.888
	(0.009)	(0.009)	(0.037)	(0.036)	(0.033)
Daily Discount Factor: δ	0.998	0.997	0.999	1.001	1.000
	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)
Monetary Curvature Parameter: α	0.975	0.976			
	(0.006)	(0.005)			
Cost of Effort Parameter: γ			1.624	1.557	1.589
			(0.114)	(0.099)	(0.104)
# Observations	1500	1125	800	800	1600
# Clusters	75	75	80	80	80
Job Effects					Yes
$H_0 : \beta = 1$	$\chi^2(1) = 8.77$	$\chi^2(1) = 1.96$	$\chi^2(1) = 7.36$	$\chi^2(1) = 11.43$	$\chi^2(1) = 11.42$
	$(p < 0.01)$	$(p = 0.16)$	$(p < 0.01)$	$(p < 0.01)$	$(p < 0.01)$
$H_0 : \beta(Col.\ 1) = \beta(Col.\ 5)$	$\chi^2(1) = 6.37$				
	$(p = 0.01)$				
$H_0 : \beta(Col.\ 2) = \beta(Col.\ 5)$		$\chi^2(1) = 8.26$			
		$(p < 0.01)$			

Notes: Parameters identified from two-limit Tobit regressions of equations (6) and (4) for monetary discounting and effort discounting, respectively. Parameters recovered via non-linear combinations of regression coefficients. Standard errors clustered at individual level reported in parentheses, recovered via the delta method. Effort regressions control for Job Effects (Task 1 vs. Task 2). We use Chi-squared tests for the null hypotheses in the last three rows.

of dynamic consistency, $H_0 : \beta = 1$: $\chi^2(1) = 1.96$, $(p = 0.16)$. These estimates indicate very limited present bias in monetary discounting.

In both specifications, we estimate α of around 0.975 indicating limited utility function curvature over monetary payments. Finding limited curvature over money is important in its own right, as linear preferences over monetary payments is indicative of fungibility. There is no desire to smooth monetary payments as there might be for consumption, with subjects treating money received at different points in time effectively as perfect substitutes. Supporting these estimates, we note that 86% of monetary allocations are corner solutions and 61% of subjects have zero interior allocations in twenty decisions.

Our non-parametric and parametric results closely mirror the aggregate findings of Andreoni and Sprenger (2012a) and Gine et al. (2010).[30] A potential concern of these earlier studies that carefully control transaction costs and payment reliability, is that a payment in the present was implemented by a payment in the afternoon of the same day, e.g. by 5:00 pm in the subjects' residence mailboxes in Andreoni and Sprenger (2012a). In this paper, because subjects repeatedly have to come to the lab, a payment in the present is implemented by an immediate cash payment. The fact that we replicate the earlier studies that carefully control for transaction costs and payment reliability alleviates the concerns that payments in the afternoon are not treated as present payments.

We confirm the finding of limited present bias in the domain of money. This could be either because the good in question, money, is fungible (a hypothesis for which we find some evidence). Alternatively, it could be because present bias in the form provided by models of dynamic inconsistency does not exist or exists in only very limited form. This motivates our exploration of choices over effort, which we believe is closer to consumption than money is.

4.2 Effort Discounting

Subjects make a total of 40 allocation decisions over effort in our seven week experiment. Twenty of these decisions are made in the first block of the experiment, and twenty in the second block. One focus of our design is testing whether participants identified as being present-biased in Block 1 demand commitment in Block 2. Hence, we opt to present here allocation data from only the first block of the experiment. This allows the prediction of commitment demand to be conducted truly as an out-of-sample exercise. In Appendix B.3 we present results of present bias from both blocks of the experiment and document very similar results.

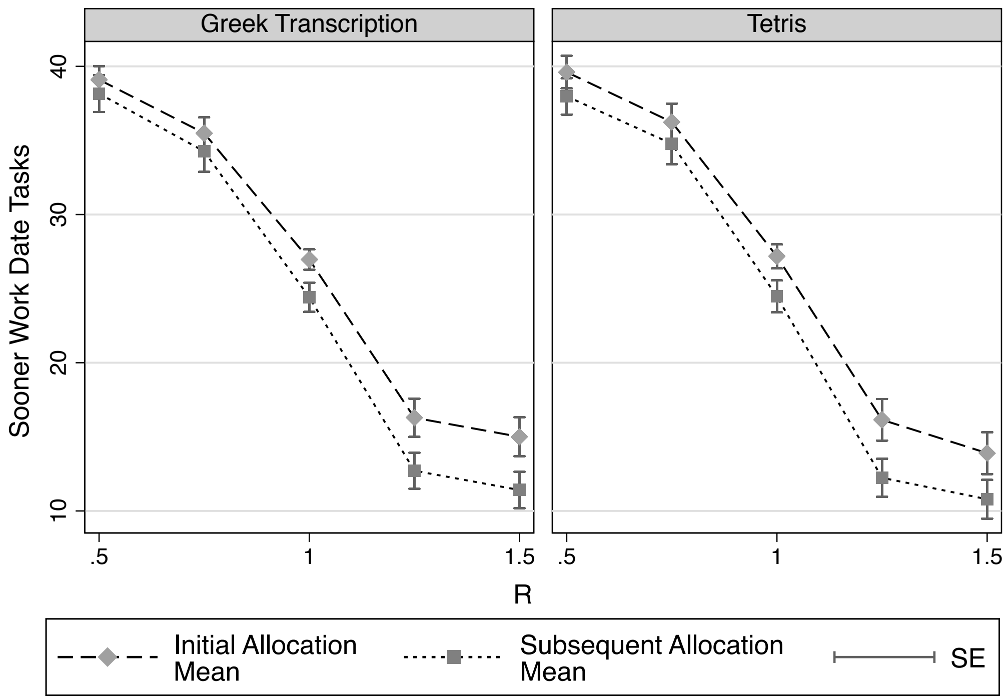
In Figure 5, we show for each value of R, the amount of tasks allocated to the sooner work date, Week 2, which could range from 0 to 50.[31] We contrast initial allocations of effort made in

[30] In both of these prior exercises substantial heterogeneity in behavior is uncovered. In subsection 4.3 we conduct individual analyses, revealing similar findings.

[31] The data are presented with the R, as opposed to relative price, to provide a standard downward sloping demand curve. When R is low, sooner tasks are relatively cheap to complete, and when R is high, sooner tasks

Week 1 with subsequent allocations made in Week 2 for the 80 subjects of the primary sample. Standard errors bars are clustered at the individual level.

Figure 5: Real Effort Discounting Behavior



As with monetary discounting, subjects appear to have understood the central intertemporal tradeoffs of the experiment as both initial and subsequent allocations decrease as R is increased. At the individual level, 95% of choices are monotonically decreasing in R, and only 5 subjects exhibit more than 5 non-monotonicities in their effort choices.[32] This suggests that subjects as a whole understand the implied intertemporal tradeoffs and the decision environment.

are relatively expensive to complete.

[32]Subjects have 32 opportunities to violate monotonicity comparing two adjacent values of R in their 40 total CTB choices. 41 of 80 subjects are fully consistent with monotonicity and only 5 subjects have more than 5 non-monotonicities. Deviations are in general small with a median required allocation change of 3 tasks to bring the data in line with monotonicity. Three subjects have more than 10 non-monotonicities indicating upward sloping sooner effort curves. Such subjects may find the tasks enjoyable such that they prefer to do more tasks sooner to fewer tasks later. We believe the increased volume of non-downward sloping behavior in effort relative to money has several sources. Subjects may actually enjoy the tasks, they make more choices for effort than for money, and half of their allocations are completed outside of the controlled lab environment. Importantly, non-monotonicities decrease with experience such that in the second block of the experiment 97 percent of choices satisfy monotonicity while in the first block, only 93 percent do so, $F(1, 79) = 8.34$ ($p < 0.01$).

Apparent from the observed choices is that at all values of R average subsequent allocations lie below average initial allocations. Controlling for all R and task interactions, subjects allocate 2.47 fewer tasks to the sooner work date when the sooner work date is the present $F(1, 79) = 14.78$, $(p < 0.01)$. Subjects initially allocate 9.3% more tasks to the sooner work date than they subsequently allocate (26.59 initial vs. 24.12 subsequent).[33]

Motivated by our non-parametric analysis we proceed to estimate intertemporal parameters. Table 2 columns (3) through (5) present two-limit Tobit regressions based on (4). In column (3) the analyzed data are the allocations for Job 1, Greek Transcription. We find an estimated cost parameter $\gamma = 1.624$ (0.114). Abstracting from discounting, a subject with this parameter would be indifferent between completing all 50 tasks on one work date and completing 32 tasks on *both* work dates.[34] This suggests non-fungibility in the allocation of tasks as individuals do desire to smooth intertemporally. A further indication of non-fungibility is that in contrast to the monetary choices, only 31% of allocations are at budget corners and only 1 subject has zero interior allocations. The daily discount factor of $\delta = 0.999$ (0.004) is similar to our findings for monetary discounting.

In column (3) of Table 2 we estimate an aggregate $\beta = 0.900$ (0.037), and easily reject the null hypothesis of dynamic consistency, $\chi^2(1) = 7.36$, $(p < 0.01)$. In column (4), we obtain broadly similar conclusions for Job 2, the partial Tetris games. We aggregate over the two jobs in column (5), controlling for the job, and again document that subjects are significantly present-biased over effort.[35] The results of column (5) indicate that discount rates measured in advance of the Week 2 work date are around zero percent per week while discount rates measured on the Week 2 work date are around eleven percent per week.

[33]The behavior is more pronounced for the first block of the experiment. For both blocks combined subjects allocate 25.95 tasks to the sooner date, 1.59 more tasks than they subsequently allocate (24.38 tasks), representing a difference of around 6%, $F(1, 79) = 15.16$, $(p < 0.01)$. See Appendix B.3 for detail.

[34]In many applications in economics and experiments, quadratic cost functions are assumed for tractability and our analysis suggests that at least in our domain this assumption would not be too inaccurate.

[35]For robustness, we run regressions similar to column (5) separately for each week and note that though the cost function does change somewhat from week to week, present bias is still significantly identified as individuals are significantly less patient in their subsequent allocation decisions compared to their initial allocation decisions. Appendix Table A3 provides estimates.

Finally, our implemented analysis allows us to compare present bias across effort and money with χ^2 tests based on seemingly unrelated estimation techniques. We reject the null hypothesis that the β identified in column (5) over effort is equal to that identified for monetary discounting in column (1), $\chi^2(1) = 6.37,\ (p = 0.01)$, or column (2), $\chi^2(1) = 8.26,\ (p < 0.01)$. Subjects are significantly more present-biased over effort than over money.[36]

4.3 Individual Analysis

On aggregate, we find that subjects are significantly more present-biased over work than over money. In this sub-section we investigate behavior at the individual level to understand the extent to which present bias over effort and money is correlated within individual.

In order to investigate individual level discounting parameters we run fixed effect versions of the regressions provided in columns (2) and (5) of Table 2.[37] As discussed in section 3, we identify discounting parameters at the individual level assuming no heterogeneity in cost or utility function curvature. Individual parameter estimates of β_e, present bias for effort, and β_m, present bias for money, are recovered as non-linear combinations of regression coefficients as described in section 3.

One technical constraint prevents us from estimating individual discounting parameters with two-limit Tobit as in the aggregate analysis. In order for parameters to be estimable at the individual level with two-limit Tobit, some interior allocations are required. As noted above, 86% of monetary allocations are at budget corners and 61% of the sample has zero interior allocations. For effort discounting, 31% of allocations are at budget corners and 1 subject has zero interior allocations. To estimate individual-level discounting, we therefore use ordinary least squares for both money and effort.[38]

[36]In Appendix B.3 we conduct identical analysis using both Blocks 1 and 2 and arrive at the same conclusions. See Appendix Table A4 for estimates.

[37]We choose to use the measures of present bias based on three week delay choices for the monetary discounting for continuity with our non-parametric tests of present bias. Further, when validating our individual measures, we focus on allocations over three week delay decisions as in the presentation for the aggregate data. Very similar results are obtained if we use the fixed effects versions of Table 2, column (1).

[38]Nearly identical aggregate discounting estimates are generated when conducting ordinary least squares versions of Table 2. Curvature estimates, however, are sensitive to estimation techniques that do and do not

Figure 6 presents individual estimates and their correlation. First, note that nearly 60% of subjects have an estimated β_m close to 1, indicating dynamic consistency for monetary discounting choices. This is in contrast to only around 25% of subjects with β_e close to 1. The mean value for β_m is 0.99 ($s.d. = 0.06$), while the mean value for β_e is 0.91 ($s.d. = 0.20$). The difference between these measures is significant, $t = 3.09$, ($p < 0.01$). Second, note that for the majority of subjects when they deviate from dynamic consistency in effort, they deviate in the direction of present bias.

Since correlational studies (e.g., Ashraf et al., 2006; Meier and Sprenger, 2010) often use binary measures of present bias, we define the variables 'Present-Biased'$_e$ and 'Present-Biased'$_m$ which take the value 1 if the corresponding estimate of β lies strictly below 0.99 and zero otherwise. We find that 56% of subjects have a 'Present-Biased'$_e$ of 1 while only 33% of subjects have a 'Present-Biased'$_m$ of 1. The difference in proportions of individuals classified as present-biased over work and money is significant, $z = 2.31$, ($p = 0.02$).[39]

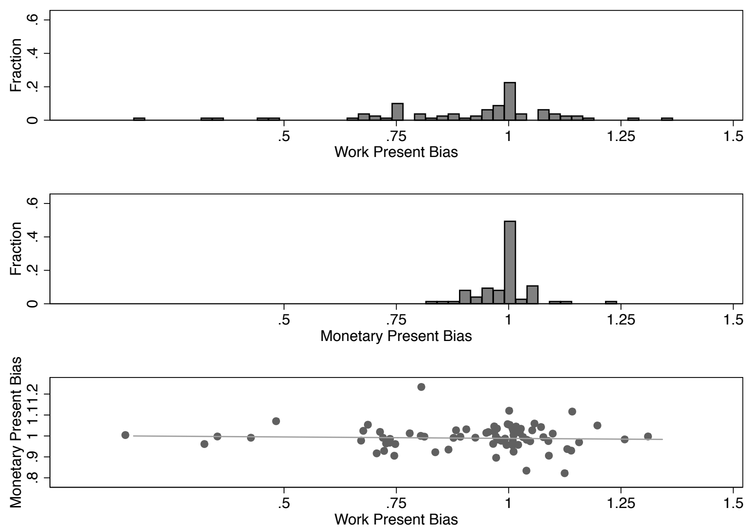
Two important questions with respect to our individual measures arise. First, how much do these measures correlate within individual? The answer to this question is important for understanding both the validity of studies relying on monetary measures and the potential consistency of preferences across domains. Significant correlations would suggest that there may be some important preference-related behavior uncovered in monetary discounting studies.[40] Figure 6 presents a scatterplot of β_m and β_e. In our sample of 75 subjects with both complete monetary and effort discounting choices, we find that β_e and β_m have almost zero correlation,

recognize that the tangency conditions implied by (4) and (6) may be met with inequality at budget corners. Se Andreoni and Sprenger (2012a) for further discussion.

[39]Further, one can define future bias in a similar way. 17% of subjects are future biased in money while 29% of subjects are future biased over effort. Similar differing proportions between present and future bias have been previously documented (see, e.g., Ashraf et al., 2006; Meier and Sprenger, 2010). Two important counter-examples are Gine et al. (2010) who find almost equal proportions of present and future biased choices and Dohmen, Falk, Huffman and Sunde (2006) who find a greater proportion of future-biased than present-biased subjects.

[40]Indeed psychology provides some grounds for such views as money generates broadly similar rewards-related neural patterns as more primary incentives (Knutson, Adams, Fong and Hommer, 2001), and in the domain of discounting evidence suggests that discounting over primary rewards, such as juice, produces similar neural images to discounting over monetary rewards (McClure, Laibson, Loewenstein and Cohen, 2004; McClure et al., 2007).

Figure 6: Individual Estimates of Present Bias



$\rho = -0.05,\ (p = 0.66)$. Additionally, we find that the binary measures for present bias, 'Present-Biased'$_e$ and 'Present-Biased'$_m$ are also uncorrelated, $\rho = 0.11,\ (p = 0.33)$.[41]

The second question concerning our estimated parameters is whether they can be validated in sample. That is, given that β_e and β_m are recovered as non-linear combinations of regression coefficients, to what extent do these measures predict present-biased allocations of tasks and money? In order to examine this internal validity question, we generate distance measures for allocations. For effort choices we calculate the budget share of each allocation for Week 2 effort. The difference in budget shares between subsequent allocation and initial allocation is what we term a 'Budget Share Distance.'[42] As budget shares are valued between $[0, 1]$, our budget share distance measure takes values on the interval $[-1, 1]$, with negative numbers indicating present-biased behavior. Each subject has 10 such effort budget share distance measures in

[41] Interestingly, when using both Blocks 1 and 2 of the data, we come to a slightly different conclusion. Though β_m and β_e remain virtually uncorrelated, with the additional data we uncover a substantial and significant correlation between Present-Biased'$_e$ and 'Present-Biased'$_m$ $\rho = 0.24,\ (p = 0.03)$. Further, 'Present-Biased'$_m$ is also significantly correlated with the continuous measure β_e, $\rho = -0.27,\ (p = 0.02)$. More work is needed to understand the relationship between monetary and effort present bias parameters.

[42] Specifically, given an initial Week 1 allocation of e_2 of work to be done in Week 2 and a subsequent allocation of e_2' in Week 2 of work to be done in week 2, the budget share distance is $\frac{e_2' - e_2}{m}$.

Table 3: Validation of Individual Parameter Estimates

Dependent Variable:	*Budget Share Distance*			
	Effort Discounting		Monetary Discounting	
	(1)	(2)	(3)	(4)
Real Effort Present Bias Parameter: β_e	0.532***			
	(0.053)			
Present-Biased$_e$ (=1)		-0.123***		
		(0.020)		
Monetary Present Bias Parameter:β_m			2.393***	
			(0.052)	
Present-Biased$_m$ (=1)				-0.201***
				(0.026)
Constant	-0.531***	0.020***	-2.391***	0.044***
	(0.052)	(0.007)	(0.049)	(0.015)
Job Effects	Yes	Yes	-	-
Choice Set Effects	-	-	Yes	Yes
# Observations	800	800	750	750
# Uncensored Observations	798	798	731	731
# Clusters	80	80	75	75

Notes: Coefficients from tobit regressions of budget share distance $\in [-1, 1]$ on identified individual discounting parameters. 10 allocations per individual for effort decisions and 10 allocations per individual for monetary decisions. Standard errors clustered on individual level in parentheses. Job fixed effects for effort and choice set fixed effects for monetary discounting included but not reported. Discounting parameters identified from OLS regressions for monetary discounting and real effort discounting with individual specific effects for both δ and β. Curvature parameter, α, and cost parameter, λ, assumed constant across individuals. Effort regressions identifying parameters control for Job Effects (Job 1 vs. Job 2). Monetary Present Bias (=1) and Effort Present Bias (=1) calculated as individuals with estimated $\beta < 0.99$ in the relevant domain. Levels of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Block 1. A similar measure is constructed for monetary discounting choices. Taking only the three week delay data, at each value of P we take the difference between the future allocation (Week 4 vs. Week 7 (Prospective)) budget share and the present allocation (Week 1 vs. Week 4 or Week 4 vs. Week 7) budget share. This measure takes values on the interval $[-1, 1]$, with negative numbers indicating present-biased behavior. Each subject has 10 such monetary budget share distance measures.[43]

Table 3 presents a validation table with tobit regressions of 'Budget Share Distance' for effort and money on our corresponding parameter estimates.[44] Individuals who are identified as present-biased by our parameter measures do indeed have more present-biased behavior for both work and money. Columns (1) and (3) demonstrate that individuals with lower values of β_e and β_m make more present-biased allocations, and those subjects with $\beta = 1$ would be predicted to make virtually identical allocations through time. Columns (2) and (4) demonstrate that subjects who are present-biased over effort allocate 12 percent less of their budget to the sooner date when the sooner date is the present. Subjects who are present-biased over money allocate around 20 percent less of their budget to the sooner payment when the sooner payment is in the future. The internal validity documented in Table 3 gives us confidence that our parameter estimates for present bias are indeed tightly linked with present-biased data patterns, appropriately capturing the behavior.

In the next section we move out-of-sample to investigate commitment demand. The investigation of commitment demand is critical to ruling out potential alternative explanations for time inconsistency in effort allocations. Our preferred explanation is the existence of a present bias in individual decision-making. Many alternative explanations exist, which are considered in detail in Section 5. Importantly, we will show that under none of these alternative hypotheses would we expect a clear link between the behavioral pattern of reallocating fewer tasks to the present and commitment demand. Further, we note that time consistent subjects that may

[43] The budget share distances for effort and money can also be used to demonstrate the extent of dynamic inconsistency. The average budget share distance for effort is -0.0494 (clustered s.e. = 0.0128) indicating that subjects allocate around 5% less of their work budget to sooner work date when allocating in the present. This value deviates significantly from the dynamically consistent benchmark of 0, $F(1, 79) = 14.87$, $(p < 0.01)$. At the individual level, 49 of 80 subjects have an average budget share distance of less than zero, 13 have an average distance of exactly zero, and 18 have an average distance greater than zero, demonstrating a modal pattern of present bias. For money, the average budget share distance is -0.0288 (0.0154), which differs marginally significantly from the dynamically consistent benchmark of 0, $F(1, 74) = 3.50$, $(p = 0.07)$. At the individual level, 28 of 75 subjects have an average budget share distance of less than zero, 32 have an average distance of exactly zero, and 15 have an average distance greater than zero, demonstrating a modal pattern of dynamic consistency. These non-parametric data closely echo the parametric findings of Figure 6 which shows substantial present bias in effort and limited present bias in money.

[44] Tobit regressions are implemented to account for the dependent variable being measured on the interval $[-1, 1]$.

be exposed to changing cost functions may actually demand flexibility. This is in contrast to a model of present bias under the assumption of sophistication. Sophisticated present-biased individuals may have demand for commitment. In the next section we document commitment demand on the aggregate level and link commitment choice to measured present-bias at both the aggregate and individual level.

4.4 Commitment

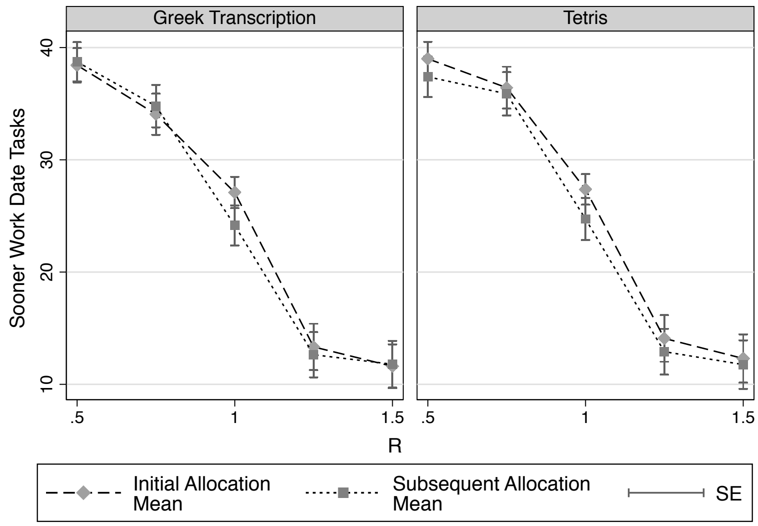
In Week 4 of our experiment, subjects are offered a probabilistic commitment device. Subjects are asked whether they prefer the allocation-that-counts to come from their Week 4 allocations with probability 0.1 (plus an amount \$X) or with probability 0.9 (plus an amount \$Y), with either \$X=0 or \$Y=0. The second of these choices represents commitment and \$X - \$Y is the price of commitment.[45] We begin by analyzing the simple choice between commitment and flexibility at price zero (\$X=0 and \$Y=0) and in section 4.4.1 we explore the value of commitment. In the simple choice of commitment, 59% (47/80) of subjects choose to commit. We define the binary variable 'Commit (=1)' which takes the value 1 if a subject chooses to commit in this decision.

Figure 7 presents Block 1 task allocation behavior separated by commitment choice in Block 2. Immediately apparent from Figure 7 is that experimental behavior separates along commitment choice. Subjects who choose commitment in Week 4 made substantially present-biased task allocations in Week 2 given their initial Week 1 allocations. Controlling for all task rate and task interactions, subjects who choose commitment allocate 3.58 fewer tasks to the sooner work date when it is the present, $F(1, 46) = 12.18$, ($p < 0.01$). Subjects who do not demand commitment make more similar initial allocations and subsequent allocations of effort. Controlling for all task rate and task interactions, they only allocate 0.89 fewer tasks to the sooner work date when it is the present, $F(1, 32) = 4.01$, ($p = 0.05$). Furthermore, subjects

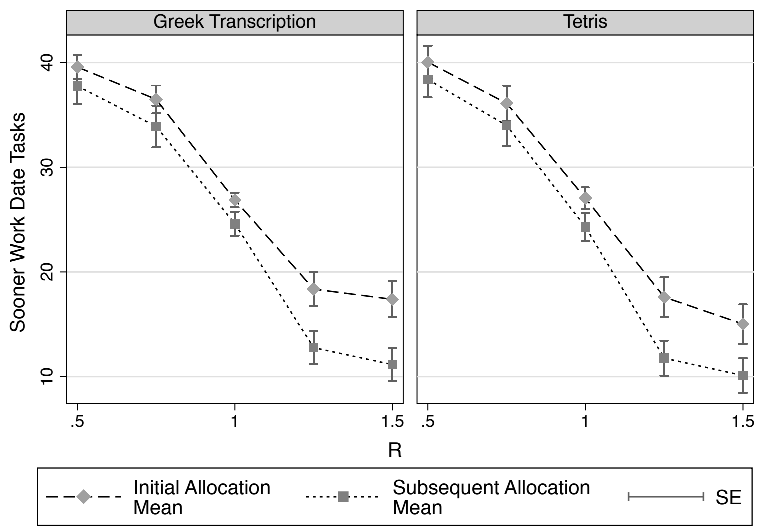
[45]To avoid cutting the sample further, here we consider all 80 subjects in the primary sample. 4 of 80 subjects switched multiple times in the commitment device price list elicitation. Identical results are obtained excluding such individuals.

Figure 7: Commitment Choice and Allocation Behavior

Panel A: Commit (=0)



Panel B: Commit (=1)



who demand commitment in Week 4 altered their allocations by significantly more tasks than subjects who did not demand commitment, $F(1, 79) = 5.84,\ (p = 0.02)$.

Table 4 generates a similar conclusion with parametric estimates. In columns (3) and (4), we find that subjects who choose commitment in Block 2 are significantly present-biased over effort in Block 1, $\chi^2(1) = 9.00,\ (p < 0.01)$. For subjects who do not choose commitment, we cannot reject the null hypothesis of $\beta = 1$ at conventional levels, $\chi^2(1) = 2.64,\ (p = 0.10)$. Further, we reject the null hypothesis of equal present bias across committers and non-committers, $\chi^2(1) = 4.85,\ (p = 0.03)$.[46]

In columns (1) and (2) of Table 4 we repeat this exercise, predicting commitment choice for effort using present bias parameters from monetary decisions. While subjects who demand commitment also seem directionally more present-biased for monetary decisions than subjects who do not demand commitment, the difference is not significant, $(p = 0.26)$.

In Table 5 we assess whether present bias identified at the individual level predicts commitment demand. We show logit regressions with 'Commit(=1)' as the dependent variable and measures of present bias over work and money as independent variables. In column (1) we find that the continuous value β_e significantly predicts commitment demand. Column (2) shows that while the binary measure, Present-Biased$_e$, has the right sign, it fails to be significant.[47] Measures of monetary present bias deliver suggestive, but statistically insignificant results in columns (3) and (4). In columns (5) and (6) we combine present bias measures and note that where significant relations are achieved, present bias over effort has substantially more predictive power than present bias over money for explaining commitment demand.

These findings indicate that present bias in effort is significantly related to future commitment choice. Individuals who are present-biased over effort are substantially more likely to choose commitment at price zero. An important caveat for this exercise is the relatively low pseudo R-squared values presented in Table 5. Though significant correlations between present bias and commitment demand are achieved, substantial variance in the choice of commitment remains unexplained.

[46] These results are stronger for the first block of the experiment prior to the offering of the commitment device, though the general patterns holds when we use both blocks of data. Appendix Table A6 provides analysis including the data from both blocks.

[47] In Appendix Table A7, we use parameter estimates from Blocks 1 and 2 combined and find that both the continuous and the binary measure are predictive. Indeed, marginal effects indicate that individuals who are present-biased are 33 percentage points more likely to demand commitment, an increase of 56% from the mean. That binary present bias identified from the combined data set has increased predictive power may be related to the arbitrary cutoff ($\beta_e < 0.99$) employed for the measure. For example, identifying binary present bias with lower values of β_e yields significance with only Block 1 data.

Table 4: Monetary and Real Effort Discounting by Commitment

	Monetary Discounting		Effort Discounting	
	Commit (=0)	Commit (=1)	Commit (=0)	Commit (=1)
	(1) Tobit	(2) Tobit	(3) Tobit	(4) Tobit
Present Bias Parameter: β	0.999 (0.010)	0.981 (0.013)	0.965 (0.022)	0.835 (0.055)
Daily Discount Factor: δ	0.997 (0.000)	0.997 (0.001)	0.988 (0.005)	1.009 (0.005)
Monetary Curvature Parameter: α	0.981 (0.009)	0.973 (0.007)		
Cost of Effort Parameter: γ			1.553 (0.165)	1.616 (0.134)
# Observations	420	705	660	940
# Clusters	28	47	33	47
Job Effects	-	-	Yes	Yes
$H_0 : \beta = 1$	$\chi_2(1) = 0.01$ ($p = 0.94$)	$\chi_2(1) = 2.15$ ($p = 0.14$)	$\chi_2(1) = 2.64$ ($p = 0.10$)	$\chi_2(1) = 9.00$ ($p < 0.01$)
$H_0 : \beta(Col.\ 1) = \beta(Col.\ 2)$	$\chi_2(1) = 1.29$ ($p = 0.26$)			
$H_0 : \beta(Col.\ 3) = \beta(Col.\ 4)$			$\chi_2(1) = 4.85$ ($p = 0.03$)	

Notes: Parameters identified from OLS regressions of equations (1) and (2) for monetary discounting and real effort discounting. Parameters recovered via non-linear combinations of regression coefficients. Standard errors clustered at individual level reported in parentheses, recovered via the delta method. Commit (=1) or Commit (=0) separates individuals into those who did (1) or those who did not (0) choose to commit at a commitment price of zero dollars. Effort regressions control for Job Effects (Job 1 vs. Job 2). Tested null hypotheses are zero present bias, $H_0 : \beta = 1$, and equality of present bias across commitment and no commitment, $H_0 : \beta(Col.\ 1) = \beta(Col.\ 2)$ and $H_0 : \beta(Col.\ 3) = \beta(Col.\ 4)$.

It is comforting for a theory of sophisticated present bias to find that present bias predicts commitment demand. However, the result is only meaningful if we can show that commitment places a binding constraint on subjects' behavior. Do individuals who demand commitment actually restrict their own activities, forcing themselves to complete more work than they

Table 5: Predicting Commitment Demand

	Dependent Variable : Commit (=1)					
	(1)	(2)	(3)	(4)	(5)	(6)
β_e	-2.595** [-0.625]				-3.157** [-0.725]	
	(1.170)				(1.333)	
Present-Biased$_e$ (=1)		0.031 [0.008]				0.301 [0.070]
		(0.459)				(0.488)
β_m			-3.146 [-0.735]		-3.841 [-0.883]	
			(4.140)		(4.008)	
Present-Biased$_m$ (=1)				0.622 [0.140]		0.588 [0.133]
				(0.533)		(0.537)
Constant	2.746**	0.336	3.635	0.323	7.262*	0.180
	(1.087)	(0.340)	(4.092)	(0.288)	(4.259)	(0.369)
# Observations	80	80	75	75	75	75
Log-Likelihood	-52.031	-54.218	-49.280	-48.838	-46.529	-48.646
Pseudo R^2	0.040	0.000	0.006	0.014	0.061	0.018
Mean of Dependent Variable	0.59	0.59	0.63	0.63	0.63	0.63

Notes: Coefficients from logistic regression of demand for commitment on identified individual discounting parameters. Marginal effects in brackets. Robust standard errors in parentheses. Commit (=1) or Commit (=0) separates individuals into those who did (1) or those who did not (0) choose to commit at a commitment price of zero dollars. Discounting parameters identified from OLS regressions of equations (1) and (2) for monetary discounting and real effort discounting with individual specific effects for both δ and β. Curvature parameter, α, and cost parameter, λ, assumed constant across individuals. Effort regressions identifying parameters control for Job Effects (Job 1 vs. Job 2). Present-Biased$_m$ (=1) and Present-Biased$_e$ (=1) calculated as individuals with estimated $\beta < 0.99$ in the relevant domain. Levels of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

instantaneously desire?[48] Given the nature of our commitment device, commitment will bind whenever initial allocations differ from subsequent allocations. Two such comparisons are considered. First, we consider the first block of the experiment when no commitment contract is available. How many more tasks would subjects have been required to complete in Week 2 had commitment been in place? To answer this question we examine budget share distances for Block 1. Non-committers have a mean budget share distance of -0.018 (clustered s.e. = 0.009) allocating about 2 percentage points less of each budget to Week 2 when deciding in

[48]Though our offered commitment contract allows individuals only to meaningfully restrict themselves, this need not be the case. One example would be to have individuals commit to completing at least 1 task at the sooner work date. As virtually all initial allocations and subsequent allocations satisfy this condition anyways, such commitment would not be meaningful and as such, should not serve as evidence for the theoretically predicted link between sophisticated present bias and commitment demand.

the present. In contrast, committers have a mean budget share distance of -0.072 (0.020), allocating 7 percentage points less to Week 2 when deciding in the present. While both values are significantly different from zero ($F(1,79) = 4.14$, ($p = 0.05$), $F(1,79) = 12.39$, ($p < 0.01$), respectively), the difference between the two is also statistically significant, $F(1,79) = 5.88$, ($p = 0.02$). Hence, had commitment been in place in Week 2 and had subjects made the same choices, committers would have been required to complete significantly more work than they instantaneously desired and would have been more restricted than non-committers. The same analysis can be done for Block 2 focusing on required work in Week 5. Non-committers have a mean budget share distance of 0.011 (0.017) while committers have a mean distance of -0.030 (0.013). The difference for committers remains significantly different from zero, $F(1,79) = 5.57$, ($p = 0.02$), and the difference between the two remains significant at the 10% level $F(1,79) = 3.68$, ($p = 0.06$).[49] Hence, in the presence of commitment in Week 5, committed subjects are required to complete significantly more work than they instantaneously desire and are more restricted than non-committed subjects.

We are aware of two prior exercises exploring the potential extent of present bias and its correlation with commitment demand. Kaur et al. (2010) link the apparently present-biased behavior of working harder on paydays with demand for a dominated wage contract wherein individuals choose a work target. If the work target is not met, an individual receives a low piece-rate wage, while if it is met or exceeded the individual receives a higher piece rate wage. As the dominated wage contract can be viewed as a commitment to complete a certain amount of work, this represents a potential link between commitment and present bias. Commitment levels are chosen by individuals themselves and are set to around one-sixth of daily production on average. Calculations indicate that committing subjects would have missed their target with probability around 0.091 in the absence of commitment, and do miss their target with commitment in place with probability 0.026. Hence, commitment can viewed as binding in about 7.5 percent of cases, effectively forcing an individual to do more work than they instantaneously desire.

[49]The difference for non-committers is no longer significantly different from zero $F(1,79) = 0.39$, ($p = 0.53$).

Ashraf et al. (2006) consider hypothetical intertemporal choices over money, rice and ice cream and link those to take-up of a savings commitment device. The authors show that present bias in the hypothetical monetary decisions is significantly correlated at the 10% level with take-up for women.

We contrast two dimensions of our study with these prior findings. The first concerns the techniques used to measure dynamic inconsistency, and the second is the extent to which subjects are bound by commitment. As opposed to monetary discounting measures or dynamic inconsistency inferred from payday effects, we attempt to measure discounting directly with intertemporal allocations of effort delivering identification. As opposed to commitments with somewhat limited binding probabilities, our committing subjects are clearly bound by commitment.

4.4.1 The Value of Commitment

A natural question is how much should subjects be willing to pay for commitment. In our theoretical section we present the value of commitment, V, as the utility difference between the discounted costs of commitment and flexibility. Given our experimental structure we can only assess the monetary value of commitment. Virtually nobody is willing to pay more than \$0.25 for commitment, with 91 percent of subjects preferring flexibility when the cost of commitment (\$X - \$Y) is \$0.25. Likewise, nobody is willing to pay more than \$0.25 for flexibility, with 90 percent of subjects preferring commitment when the price of commitment (\$X - \$Y) is -\$0.25. Taking the midpoint of each person's price list switching interval, the data thus imply a median valuation of \$0.125.[50] For committers the median valuation is \$0.125 while for non-committers the median valuation is \$-0.125.

What do these monetary valuations imply for the extent of V and correspondingly for the extent of sophistication? That present bias is predictive of commitment demand at price zero indicates at least partial sophistication on average, $\widehat{\beta} < 1$. Further analysis requires a

[50]For this measure we exclude the four individuals with multiple switching.

transformation between the value of commitment, V, and its monetary valuation, along with explicit assumptions on the benefits of flexibility.

We begin by recalling the valuation of commitment for stationary cost functions,

$$V = (1-2p)\cdot\beta\delta^{t-s}\cdot\{[(e^{s*}_{t,t}+\omega)^{\gamma}+\delta^{k}c(e^{s*}_{t+k,t}+\omega)^{\gamma}]-[(e^{*}_{t,s}+\omega)^{\gamma}+\delta^{k}c(e^{*}_{t+k,s}+\omega)^{\gamma}]\},$$

and note that V can be calculated directly for the fully sophisticated benchmark of $\widehat{\beta}=\beta$, which implies $(e^{s*}_{t,t}, e^{s*}_{t+k,t}) = (e^{*}_{t,t}, e^{*}_{t+k,t})$. For example, using the parameters estimates of Table 4, columns (3) and (4) and the actual allocations at $R=1$, we can calculate the fully-sophisticated value of commitment for committing and non-comitting subjects. For comitting subjects, we calculate $V_{C=1} = 1.23$, which can be expressed in equivalent number of tasks as $c^{-1}(1.23) = 1.14$ tasks. For non-comitting subjects, we calculate $V_{C=0} = -2.06$, which can be expressed in equivalent number of tasks as -1.59 tasks.

Comparing these valuations and equivalent tasks to monetary measures requires a set of assumptions for translating the discounted costs of tasks into dollars. At the average minimum work completion rate ($\sim$ 1 task per minute), subjects can complete approximately 60 tasks per hour. Assuming earnings of around \$12 per hour and a constant task value, a subject would be willing to complete 1 task for around \$0.20.[51] Hence the monetary value of commitment should be around \$0.23 for committing subjects and the value of flexibility should be around \$0.32 for non-committing subjects. These values compare favorably to the monetary valuations reported above. Hence, assuming complete sophistication and no additional benefits to flexibility, we predict monetary commitment valuations reasonably close to the valuations expressed by subjects.

The provided exercise provides some initial intuition on the extent of sophistication. If individuals are fully sophisticated, monetary valuations for commitment should be close to

[51]The assumption of constant per task reservation value is important. With convex costs an individual should have a lower reservation value for the first task than the sixtieth. We opt to present the average valuation recognizing the possibility that valuations could be either higher or lower. Appendix C analyzes the value of commitment demand at a wide range of potential per task valuations to provide sensitivity analysis.

those observed. Naturally, evaluating $\hat{\beta} > \beta$ lowers the value of commitment and for $\hat{\beta} = 1$ commitment should be worth exactly zero. In Appendix C we analyze specific values of $\hat{\beta}$ and corresponding valuations for commitment under various assumptions for the transformation of V to dollars. This analysis also considers all allocations, not only those at one interest rate. Clear from this exercise is that under the assumption of no additional benefits to flexibility, only in extreme cases should commitment be worth more than a dollar.

We are hesitant to draw strong conclusions beyond the plausibility of a high degree of sophistication from our commitment valuation data. First, given the ex-post parameter estimates, our elicitation procedure clearly was not optimized for fine price differentiations. Second, it is possible that subjects largely followed the money in the elicitation, preferring either commitment or flexibility depending on which option provided additional payment. A direct experiment precisely identifying $\widehat{\beta}$ is a clear next step that research in this vein should take.

4.5 Between Subjects Replication Exercise

A key contribution of our data is the documentation of limited present bias in the domain of money and more substantial present bias in the domain of work. One potential interpretation for this data is that models of dynamic inconsistency are validated when tested in their relevant domain and that time dated choices over fungible monetary payments cannot easily speak to the predictions of such models.

However, in the presented within-subjects study, several design choices were made that might muddy this interpretation. First, subjects faced different interest rates and forms of budget constraint for effort and for money.[52] Second, the delay lengths for money were three to six weeks, while the delay lengths for effort were only one week. Third, subjects always completed their effort allocations prior to completing their monetary allocations. Fourth, present bias is identified for effort from only a dynamic choice, while present bias is identified for

[52] That is, the constraint for effort was of a present value form, $e_t + Re_{t+k} = V$, while the constraint for money was of a future value form, $Rc_t + c_{t+k} = m$.

money from a combination of static and dynamic choices.[53] Fifth, for effort one allocation was chosen to be the allocation-that-counts from the initial and subsequent allocations with an asymmetric probability, while for money each allocation could be the alllocation-that-counts with equal probability. Further, the Week 4 monetary choices were paid separately from the Week 1 choices. Though each design choice has a natural motivation, including our desire to replicate prior exercises, one could potentially imagine them influencing the degree of dynamic inconsistency.[54]

To alleviate these concerns, we conducted a between subjects replication exercise. 200 subjects, again from the UC Berkeley Xlab subject pool, were randomized into two conditions: one in which allocations were made for money and one in which allocations were made for greek transcription. In both conditions subjects selected into a four week study on decision-making over time and were informed that their earnings would be approximately $60 if all aspects of the study were completed. The main goal of the replication exercise is to keep allocation decisions identical, with the only difference being whether allocations are over money or effort.

Mirroring our effort study, in Week 1 of the replication exercise subjects make allocations over Weeks 2 and 3. In Week 2, subjects again make allocations over Weeks 2 and 3. All allocations are made on a study website either in the lab in Week 1 or on any computer with internet access in Week 2. In Week 2, one of the Week 1 or Week 2 decisions is chosen at random, with each having equal probability, and the corresponding allocation is implemented.

[53] That is, for effort to identify present bias one compares the Week 1 allocations over Weeks 2 and 3 to the Week 2 choices over Weeks 2 and 3. For money to identify present bias one compares the Week 1 allocations over Weeks 4 and 7 to the Week 4 choices over Weeks 4 and 7, the Week 1 allocations over Weeks 1 and 4 to the Week 1 allocations over Weeks 4 and 7, and the Week 1 allocations over Weeks 1 and 4 to the Week 1 allocations over Weeks 1 and 7.

[54] The specific rationale for each choice, respectively: first, we expected substantially more curvature for effort than money, which suggests different interest rates to avoid corner solutions. Second, we organized the monetary choices around dates the subjects would come to the lab to equalize transactions costs. Third, our primary focus was the effort choices, hence we sought to ensure theses data were collected. Fourth, we wished to replicate the standard static evidence on present bias in money and benefited from an opportunity in Week 4 to additionally generate dynamic evidence. Fifth and sixth, we did not wish to burden the subjects with another, potentially complicated, procedure for determining which monetary decision would be implemented.

For both effort and money, allocations are made using budgets of the form,

$$Ra_2 + a_3 = m.$$

Where a_2 refers to an allocation of either effort or money to Week 2 and a_3 refers to an allocation of either effort or money to Week 3. For both effort and money $R \in \{0.66, 0.8, 0.91, 0.95, 1, 1.05, 1.11, 1.25, 1.54\}$, covering the interest rates used for both money and effort from our initial experiment. For money $m = \$20$ and for effort $m = 60$ tasks, such that units are easily matched by dividing by three. Following our prior study, minimum payments of \$5 for money and minimum work of 10 tasks for effort are implemented in Weeks 1, 2, and 3.

We attempt to put precise time stamps on both the completion of tasks and the collection of money. For effort, subjects are told they must complete their tasks from the chosen allocation on a study website between 9 am and 6 pm on the relevant day in Weeks 2 and 3. For money, subjects are told they must collect their payments from the chosen allocation at the UC Berkeley Xlab between 9 am and 6 pm on the relevant day in Weeks 2 and 3. To make the Week 2 allocations as immediate as possible, subjects are additionally told in advance they will have to either complete their Week 2 tasks or collect their Week 2 funds within two hours of making their Week 2 allocations. Appendix E has the full study instructions.

If subjects complete all aspects of the study, including collecting their money or completing their tasks on each relevant date within the relevant time window, they are eligible for a completion payment paid in the fourth week of the study. For effort, the completion payment is \$60 with a non-completion payment of \$5. For money, the completion payment is \$30 with a non-completion payment of \$5. All payments, including those from monetary allocations, are made in cash at the Xlab by a single research assistant who remained in place from 9 am to 6 pm on the relevant dates. All 200 subjects began the study on Thursday April 17, 2014. Of these a total of 194 completed the study on Thursday May 1, 2014, with 95 from the effort condition and 99 from the money condition.

In this between subjects design, direct comparisons in present bias across conditions can be made. Figure 8 plots the number of tasks in Panel A (out of 60) and amount of money in Panel B (out of \$20) allocated to Week 3 for each level of R. Separate series are provided for when the allocation is made in Week 1 and in Week 2. Note that because the budget constraints are identical, Week 3 tasks are decreasing in R, while Week 3 money is increasing in R. Note as well that due to the form of the budget constraint, it is the Week 3 units, which have constant value, that are graphed.[55]

Figure 8 closely reproduces our prior within-subject findings. At each value of R, individuals appear present-biased for effort, allocating more effort to the later date when the sooner date is the present. Controlling for R, subjects allocate 2.14 (clustered s.e. = 1.10) more tasks to Week 3 in Week 2 relative to Week 1, $F(1, 94) = 3.82,\ p = 0.05$. In contrast, for money subjects appear almost perfectly dynamically consistent on average. Controlling for R, subjects allocate \$0.14 (clustered s.e. = 0.12) less to Week 3 in Week 2 relative to Week 1 $F(1, 98) = 1.37,\ p = 0.25$.

Non-parametric replication in hand, we now turn to estimation of aggregate utility parameters. In Table 6, we replicate the estimation exercise of Table 2 with the new between-subjects data. The parameter values and corresponding conclusions are effectively unchanged. For monetary present bias in column (1), we estimate $\beta = 0.997$ (clustered s.e. = 0.005), which compares favorably to Table 2, column (2), which estimates $\beta = 0.988$ (0.009). Similar to our within-subjects conclusion, we fail to reject the null hypothesis of dynamic consistency, $\beta = 1$, for money, $\chi^2(1) = 0.50,\ (p = 0.48)$. Interestingly, we also find quite similar discount factor and curvature estimates between Table 6, column (1) and Table 2, column (2). For effort present bias in column (2), we estimate $\beta = 0.892$ (0.056), which compares favorably to Table 2, column (3) for greek transcription where $\beta = 0.900$ (0.037). Similar to our within-subjects conclusion, we reject the null hypothesis of $\beta = 1$ for effort, $\chi^2(1) = 3.73,\ (p = 0.05)$. Again, we find quite similar estimates for the auxiliary parameters between Table 6, column (2) and

[55]This is in contrast to the prior figures where earlier tasks had constant value and were graphed.

Figure 8: Between Subjects Replication Exercise

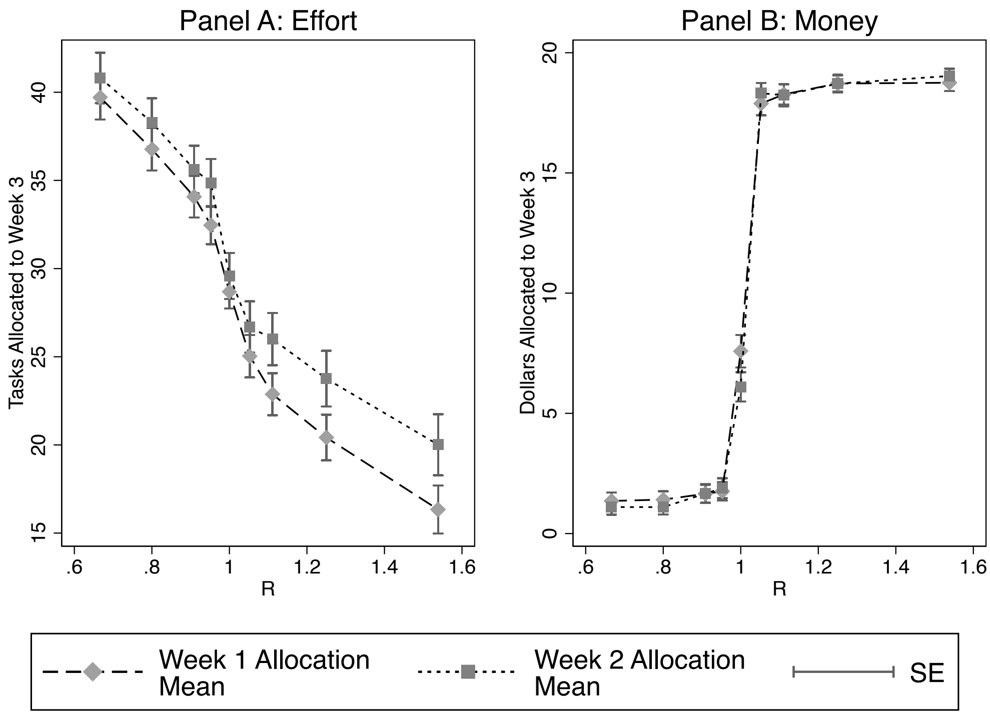


Table 2, column (3). The analysis again allows us to compare present bias across effort and money, and again we reject the null hypothesis that the β identified for money is equal to that identified for effort, $\chi^2(1) = 3.50$, $(p = 0.06)$.

Though these findings closely replicate our prior within-subjects data, it is important to note that the data from this exercise yields somewhat less precise measures and test statistics than our initial study. We hesitate to speculate as to the source of this imprecision, and draw some comfort from the replication not only of the conclusions, but also the point estimates from our prior work.

5 Discussion

Our effort discounting data address several key confounds present in monetary studies, such as fungibility and arbitrage issues. In this section we address whether we can attribute the

Table 6: Replication Exercise Parameter Estimates

	Monetary Discounting	Effort Discounting Greek
	(1)	(2))
Present Bias Parameter: β	0.997	0.892
	(0.005)	(0.056)
Daily Discount Factor: δ	0.998	1.009
	(0.001)	(0.005)
Monetary Curvature Parameter: α	0.952	
	(0.009)	
Cost of Effort Parameter: γ		1.774
		(0.167)
# Observations	1782	1710
# Clusters	99	95
$H_0 : \beta = 1$	$\chi^2(1) = 0.50$	$\chi^2(1) = 3.73$
	$(p = 0.48)$	$(p = 0.05)$
$H_0 : \beta(Col.\ 1) = \beta(Col.\ 2)$	$\chi^2(1) = 3.50$	
	$(p = 0.06)$	

Notes: Parameters identified from two-limit Tobit regressions of equations (6) and (4) for monetary discounting and effort discounting, respectively. Parameters recovered via non-linear combinations of regression coefficients. Standard errors clustered at individual level reported in parentheses, recovered via the delta method. We use Chi-squared tests for the null hypotheses in the last two rows.

observed behavior for effort choices to dynamic inconsistency. Foremost, the ability to predict commitment demand based on present-biased allocations gives a degree of confidence that present-biased allocations are driven by dynamic inconsistency. In the following, we discuss four additional hypotheses that can generate time inconsistent effort allocations. These are (unanticipated) permanent shocks to the cost function of performing the tasks, unanticipated shocks to the cost function in Week 2, general uncertainty in cost functions, and simple mistakes. Though none of these explanations would predict a correlation between time inconsistency and commitment demand, we can also address these hypotheses directly.

First, subjects may make present-biased allocations in Week 2 not because they are present-biased, but because their cost function for the tasks changed permanently. Maybe upon returning to the tasks they find them to be more or less difficult than they previously envisioned. For example, this could be because they have an injury that makes typing harder, have a bigger and better (or smaller and worse) screen at home than in the lab, which makes the tasks less (more) onerous, etc.[56] Though we do attempt to give subjects a sense of the tasks, this is a plausible and critical confound. Our environment is able to address this confound as changes to perceived cost functions are separable from time preferences. The shape of the cost function is identified from changes in the value of R. Because both initial allocations and subsequent allocations are made at various interest rates, the cost function is identified at multiple points and time. In Appendix Table A3, we estimate cost functions and discounting parameters at each point in time. We do not find evidence that cost functions change over time.[57] This lends credence to the notion that changes in cost functions are not driving the observed behavior.[58]

Second, subjects may reallocate fewer tasks to the present due to an unforeseen, local shock that resulted in an increase in the cost function in Week 2 only. This could be because the subject is unusually busy in Week 2 because of a surprise exam, or finds himself unusually exhausted and hence unusually irritated with the length of work to be done. There are several

[56]We see this channel as distinct from the role of uncertainty, as such changes in difficulty need not have been forecasted.

[57]The analysis of Appendix Table A3 can be conducted separately for committing and non-committing subjects to examine if those individuals identified to be dynamically inconsistent in their commitment choice have varying cost functions or varying discounting parameters over time. For committing subjects the daily discount factor measured in Week 1 is 1.011 (s.e. = 0.007), while the daily discount factor measured in Week 2 is 0.985 (0.006). This difference is significant at all conventional levels, $\chi^2(1) = 6.60$ ($p = 0.01$). For committing subjects the cost function parameter measured in Week 1 is $\gamma = 1.739$ (0.184), and in Week 2 is $\gamma = 1.519$ (0.121). This difference is not significant at conventional levels, $\chi^2(1) = 2.55$ ($p = 0.11$). This indicates that for subjects separately identified as present-biased through their commitment choice, changing behavior through time is more clearly linked to changing discounting parameters and not changing cost functions. No differences in either discounting or cost functions are observed between Weeks 1 and 2 for non-committing subjects.

[58]Note that if cost functions would change over time, and this were the unique driver for changes in allocations between Week 1 and Week 2, we would observe a specific pattern of allocations. If an individual moved from having an almost linear cost function to a very convex one, the corresponding allocations would shift from being very price sensitive to limited price sensitivity. When initial allocations asked for lot of work to be done in Week 2, we would indeed see a change that amounts to a reduction of work in Week 2. However, for allocations that asked for little work in Week 2, we would see an increase in work to be done in Week 2. This is not what we observe. The data show a universal reduction of work to be allocated in Week 2.

ways to address this concern. First, a simple way in which subjects may find it unusually difficult to complete the work in Week 2 is if they log on to the experimental website so late, just prior to midnight, that they have only a very limited opportunity to complete their tasks. We can check for this hypothesis because we recorded the time at which subjects made their allocations. The median subject completed their allocations in Week 2 with 10.3 hours remaining before the imposed midnight deadline. Only 4 of 80 subjects completed their allocations in Week 2 with less than 2 hours remaining before the imposed midnight deadline and 0 of 80 completed their allocations with less than 1 hour remaining. We therefore do not find evidence that a physical time constraint is a driving force in the allocations.

However, subjects logging on later may indeed be those who experienced an unanticipated shock in costs (even if their timing does not entail a physical constraint). We therefore examine whether subjects who log on to our experimental website later in the evening of their Week 2 work date exhibit more present bias. Individuals who log on with less than 4 hours before midnight (20 percent of our sample) are no more present-biased and have virtually identical allocation behavior as others.[59]

As a final way to assess whether some subjects may have had unusual shocks to their cost function (and whether these are subjects that generate our results of present-biased allocations), we can find a proxy for the costs of the tasks in Week 2. Specifically, we examine the amount of time it takes subjects to complete their minimum work in Week 2. Minimum work took the median subject around 18 minutes to complete. Those subjects who take longer than 25.7 minutes (20 percent of our sample) are no more present-biased and have virtually identical

[59]Subjects logging on with more than 4 hours before midnight allocate an average of 23.80 (s.d = 15.91) tasks to the sooner work date in Week 2, while subjects logging on with less than 4 hours allocate 25.43 (14.06). Even without accounting for multiple observations this difference is not significant, $t(798) = 1.19,\ p = 0.24$. Subjects logging on with more than 4 hours before midnight have budget share distances between Weeks 1 and 2 of -0.049 (0.21), indicating they allocate around 5 percent less of the budget of tasks to the sooner work date in Week 2 than they allocated in Week 1. Subjects logging on with less than 4 hours have budget share distances between Weeks 1 and 2 of -0.052 (0.20). Even without accounting for multiple observations this difference is not significant, $t(798) = 0.15,\ p = 0.88$. Note however, that in general, subjects that log in later may be more present-biased, as they do everything a little later. And indeed, if we instead cut at the median log-in time, 10.3 hours before midnight, marginally significant differences are observed indicating that present-biased individuals may be logging in later. However, such individuals do not appear to be those particularly close to the deadline.

allocation behavior as others.[60] Naturally, these analyses may not give a fully satisfactory response to the potential confound presented by forecasting error and boredom. If indeed such a possibility is the source of our present-biased data patterns, a final question is whether or not such a hypothesis delivers the observed correlation between present-bias and commitment demand. We believe the answer to this question to be no.

A third class of explanations which can generate a pattern of present-biased behavior in the absence of time inconsistency concerns uncertainty in cost functions. When making initial allocations, subjects do so under a different informational environment than when making their subsequent allocations. There could be uncertainty for initial allocations, which is partially resolved when allocations are again made one week later. Several aspects of uncertainty warrant attention. First, individuals may carry preferences for the resolution of uncertainty (Kreps and Porteus, 1978; Epstein and Zin, 1989; Chew and Epstein, 1989). Unlike monetary designs, in our effort experiment such a preference may more naturally lead to a future bias. Subjects desiring to resolve uncertainty in their subsequent allocation choices could, in principle, choose to complete their tasks immediately when the present is available. Second, our discounting estimates do not account for subjects' potential uncertainty on their own parameters, such as uncertainty with regards to the future costliness of the task. Though the weekly parameter estimates provided in Table A3 help to alleviate some concerns, a deeper problem may exist. Subjects may make allocations in Week 1 that minimize their discounted *expected* cost in future weeks given the potential realizations of future parameters. This uncertainty may be subsequently resolved in Week 2, such that subjects minimize their discounted cost at specific realizations of key parameters. As the minimizer of the expectation need not be the expectation of the minimizer, such issues can lead to inconsistencies between initial allocations

[60] Subjects taking less than 25.7 minutes allocate an average of 24.11 (s.d = 15.43) tasks to the sooner work date in Week 2, while subjects taking more than 25.7 minutes allocate 24.15 (16.11). Even without accounting for multiple observations this difference is not significant, $t(798) = 0.03,\ p = 0.98$. Subjects taking less than 25.7 minutes have budget share distances between Weeks 1 and 2 of -0.049 (0.20), indicating they allocate around 5 percent less of the budget of tasks to the sooner work date in Week 2 than they allocated in Week 1.Subjects taking more than 30 minutes have budget share distances between Weeks 1 and 2 of -0.053 (0.22). Even without accounting for multiple observations this difference is not significant, $t(798) = 0.23,\ p = 0.82$.

and subsequent allocations. To explore the extent to which this issue hampers identification of present bias, we conduct simulations under a variety of uncertainty structures in Appendix A. Uncertainty, unresolved at initial allocation and realized at the time of the subsequent allocation, does bias our estimates of β both at the aggregate and individual level. However, the direction of bias is generally upward in the parameter regions of interest, leading to less estimated present bias.[61] Importantly, a subject with future uncertainty would benefit from flexibility, such that even if present bias was delivered by uncertainty of some form one would not expect a correlation between present bias and commitment demand.

Fifth, present-biased allocations of effort may be a simple decision error. Hence, present bias, or any dynamic inconsistency, may be an unstable phenomenon. The two blocks of our experiment speak to this possibility. Subjects have two opportunities to exhibit present-biased allocations. Indeed, present-biased behavior in Block 1 and Block 2 is significantly correlated.[62] At the allocation level, a subject who is present-biased in Block 1 is 58% more likely than others to be present-biased in Block 2, $F(1, 79) = 6.94$, $(p = 0.010)$.[63] Additionally, an individual who is dynamically consistent in Block 1 is 85% more likely to be dynamically consistent in Block 2 than others $F(1, 79) = 50.88$, $(p < 0.01)$.[64]

This discussion helps to clarify some of the potential confounds for our observed effects.

[61] Intuitively, subjects with unresolved uncertainty on future parameters seek to avoid the extreme possibilities of working under a very convex cost structure that is only rarely realized. This leads initial allocations to be frequently lower than subsequent allocations, particularly at higher interest rates. Appendix A provides greater detail.

[62] Though the behavior is significantly correlated when examined as indicators for present bias, future bias and dynamic consistency; the budget share distances are not significantly correlated through time. This may be due to the sheer volume of data with budget share distances equal to zero and the relative lack of stability for future-biased behavior.

[63] Test statistic from OLS regression of binary indicator for a present-biased allocation in Block 2 on matched indicator for present-biased allocation in Block 1 with standard errors clustered on the subject level. The estimated constant is 0.218 ($s.e. = 0.030$) and the coefficient on Block 1 present bias is 0.128 ($s.e. = 0.049$).

[64] Test statistic from OLS regression of binary indicator for a dynamically consistent allocation in Block 2 on matched indicator for a dynamically consistent allocation in Block 1 with standard errors clustered on the subject level. The estimated constant is 0.400 ($s.e. = 0.041$) and the coefficient on Block 1 dynamic consistency is 0.342 ($s.e. = 0.048$). Interestingly, somewhat less precision is found for future biased allocations. An individual who is future-biased in Block 1 is 54% more likely to be future-biased in Block 2 than others $F(1, 79) = 3.07$, $(p = 0.08)$. Test statistic from OLS regression of binary indicator for a future-biased allocation in Block 2 on matched indicator for a future-biased allocation in Block 1 with standard errors clustered on the subject level. The estimated constant is 0.162 ($s.e. = 0.025$) and the coefficient on Block 1 future bias is 0.088 ($s.e. = 0.050$).

We view it as unlikely that present-biased allocations of effort are driven by unanticipated permanent or temporary shocks, uncertainty, or decision error. Further, that present bias over effort exhibits stability and predicts commitment demand gives confidence that our observed effects are generated by dynamic inconsistency.

6 Conclusion

Present biased time preferences are a core of behavioral research. The key hypothesis of diminishing impatience through time is able to capture a number of behavioral regularities at odds with standard exponential discounting. Further, the possibility of sophistication provides an important channel for policy improvements via the provision of commitment devices. With the exception of only a few pieces of research, most evidence of dynamic inconsistency is generated from experimental choices over time-dated monetary payments. When those are administered in a way to keep transaction costs constant and uncertainty at bay, recent studies have found limited evidence of dynamic inconsistency. However, such findings may not be appropriate to reject a model defined over streams of consumption.

The present study attempts to identify dynamic inconsistency for choices over real effort. We introduce a longitudinal design asking subjects to allocate and subsequently allocate again units of effort through time. A complementary monetary study is conducted for comparison. We document three key findings. First, in choices over monetary payments, we find limited evidence of present bias, confirming earlier work. Second, in choices over effort, we find substantial present bias. Subjects reallocate about 9% less work to the present than their initial allocation. Corresponding parameter estimates generate a similar conclusion. Individuals are estimated to be substantially present-biased in effort choices and significantly closer to dynamically consistent in choices over money. Third, we study commitment demand, documenting that at price zero roughly 60% of subjects prefer commitment to flexibility. A key result is that these commitment decisions correlate significantly with previously measured present bias. Individuals who demand commitment are significantly more present-biased in effort than those who do not. This provides validation for our experimental measures and helps to rule out a variety of potential confounds. Importantly, in our design commitment meaningfully restricts activities. Committed subjects are required to complete more effort than they instantaneously desire. By documenting the link between experimentally measured present bias and commitment demand, we provide support for models of dynamic inconsistency with sophistication. Subjects are apparently aware of their present bias and take actions to limit their future behavior.

We view our paper as providing a portable experimental method allowing tractable estimation of intertemporal preferences over consumption (effort) and correlating such preferences with a meaningful, potentially constraining, commitment device. Though the implementation here is with American undergraduates, we feel the design is suitable for field interventions.

We draw one conclusion and several words of caution from our findings. Our results indicate that present bias is plausibly identified in choices over effort and, furthermore, is linked to effort-related commitment demand. However, we caution using the estimated parameters at face value as they are for a specific subject pool (self-selected to work for six weeks for final payment in week seven) and a specific task. There may be other decision environments wherein behavior may not be well captured by models of dynamic inconsistency. For example, subjects may wish to get a painful single experience over with immediately or postpone a single pleasure (Loewenstein, 1987).[65] Lastly and most importantly, though fungibility issues may be mediated in the present design, the natural problems of arbitrage will still exist if subjects substitute effort in the lab with their extra-lab behavior. The existence and use of such substitutes, like avoiding doing laundry or homework in response to the experiment, will confound our measures in much the same way as monetary studies. Discounting will be biased towards market interest rates, present bias will be exhibited only if such rates change through time, and cost functions will be biased towards linearity. Though our data suggest effort is less fungible than money, one cannot say that extra-lab smoothing opportunities for effort are eliminated. Hence, one should view our measures as lower bounds on the true extent of dynamic inconsistency and

[65]This suggests a key anticipatory component of intertemporal behavior, potentially mediated by our design's use of minimum effort requirements and convex decisions.

the instantaneous cost of tasks. We want to, however, point out that to some extent such fungibility will be present in many dimensions in which time inconsistency has been measured. For example, people may put off laundry or homework when they go to the gym, or fill out tax forms, etc. Ultimately, the best measure of time inconsistency will be one that predicts ecologically relevant decisions across a broad set of environments. This suggests important avenues for future research.

References

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and Elisabet E. Rutstrom, "Eliciting Risk and Time Preferences," *Econometrica*, 2008, *76* (3), 583–618.

_ , _ , _ , **and** _ , "Discounting Behavior: A Reconsideration," *Working Paper*, 2012.

Andreoni, James and Charles Sprenger, "Estimating Time Preferences with Convex Budgets," *American Economic Review*, 2012, *102* (7), 3333–3356.

_ **and** _ , "Risk Preferences Are Not Time Preferences," *American Economic Review*, 2012, *102* (7), 3357–3376.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 2002, *13* (3), 219–224.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 2006, *121* (1), 635–672.

Brown, Alexander L., Zhikang Eric Chua, and Colin F. Camerer, "Learning and Visceral Temptation in Dynamic Saving Experiments," *Quarterly Journal of Economics*, 2009, *124* (1), 197–231.

Chabris, Christopher F., David Laibson, and Jonathon P. Schuldt, "Intertemporal Choice," in Steven N. Durlauf and Larry Blume, eds., *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008.

Cheung, Stephen L., "On the Elicitation of Time Preference Under Conditions of Risk," *American Economic Review*, Forthcoming.

Chew, S. H. and Larry G. Epstein, "The Structure of Preferences and Attitudes Towards the Timing of the Resolution of Uncertainty," *International Economic Review*, 1989, *30* (1), 103–117.

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

Cubitt, Robin P. and Daniel Read, "Can Intertemporal Choice Experiments Elicit Preferences for Consumption," *Experimental Economics*, 2007, *10* (4), 369–389.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Dynamic inconsistency predicts self-control problems in humans," *Working Paper*, 2006.

Epstein, Larry G. and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, *57* (4), 937–969.

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

Gine, Xavier, Jessica Goldberg, Dan Silverman, and Dean Yang, "Revising Commitments: Time Preference and Time-Inconsistency in the Field," *Working Paper*, 2010.

Halevy, Yoram, "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 2008, *98* (3), 1145–1162.

_ , "Time Consistency: Stationarity and Time Invariance," *Working Paper*, 2012.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams, "Estimating individual discount rates in Denmark: A field experiment," *American Economic Review*, 2002, *92* (5), 1606–1617.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control and the Development of Work Arrangements," *American Economic Review, Papers and Proceedings*, 2010, *100* (2), 624–628.

Knutson, Brian, Charles M. Adams, Grace W. Fong, and Daniel Hommer, "Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens," *The Journal of Neuroscience*, 2001, *21* (RC159), 1–5.

Kreps, David M. and Evan L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 1978, *46* (1), 185–200.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

Loewenstein, George F., "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 1987, *97* (387), 666–684.

McClure, Samuel, David Laibson, George Loewenstein, and Jonathan Cohen, "Separate neural systems value immediate and delayed monetary rewards," *Science*, 2004, *306*, 503–507.

— , — , — , and — , "Time discounting for primary rewards," *Journal of Neuroscience*, 2007, *27* (21), 5796–5804.

Meier, Stephan and Charles Sprenger, "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal - Applied Economics*, 2010, *2* (1), 193–210.

Miao, Bin and Songfa Zhong, "Separating Risk and Time Preferences," *Working Paper*, 2012.

O'Donoghue, Ted and Matthew Rabin, "Doing it Now or Later," *American Economic Review*, 1999, *89* (1), 103–124.

— and — , "Choice and Procrastination," *The Quarterly Journal of Economics*, 2001, *116* (1), 121–160.

Read, Daniel and Barbara van Leeuwen, "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Organizational Behavior and Human Decision Processes*, 1998, *76* (2), 189–205.

Samuelson, Paul A., "A Note on Measurement of Utility," *The Review of Economic Studies*, 1937, *4* (2), 155–161.

Solnick, Jay V., Catherine H. Kannenberg, David A. Eckerman, and Marcus B. Waller, "An Experimental Analysis of Impulsivity and Impulse Control in Humans," *Learning and Motivation*, 1980, *11*, 61–77.

Strotz, Robert H., "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 1956, *23*, 165–180.

Thaler, Richard H., "Some Empircal Evidence on Dynamic Inconsistency," *Economics Letters*, 1981, pp. 201–207.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, MA, 2002.