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: April 10, 2015

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.[2] 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.[3] Furthermore, monetary payments may not

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

[2]Recent efforts using time dated monetary payments to identify time preferences include Ashraf, Karlan and Yin (2006), Andersen, Harrison, Lau and Rutstrom (2008), Dohmen, Falk, Huffman and Sunde (2010), Tanaka, Camerer and Nguyen (2010), Benjamin, Choi and Strickland (2010) Voors, Nillesen, Verwimp, Bulte, Lensink and van Soest (2012), Bauer, Chytilova and Morduch (2012), Sutter, Kocher, Glatzle-Ruetzler and Trautmann (2013), and Dupas and Robinson (2013).

[3]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

be suitable to identify parameters of models defined over time-dated consumption. Arbitrage arguments imply that choices over monetary payments should only reveal subjects' borrowing and lending opportunities (Cubitt and Read, 2007).[4] 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

risk preferences deliver behavior observationally equivalent to present bias as described above (Andreoni and Sprenger, 2012b).

[4]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.

for deadlines for classroom and work assignments, a potential sign of commitment demand for dynamically inconsistent individuals. Though suggestive, neither exercise allows for precise identification of discounting parameters, nor delivers the critical linkage between present bias and commitment demand. With the exception of Ashraf et al. (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 3.4. Most important is the measurement of dynamic inconsistency. As opposed to monetary measures or measuring 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 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, 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 second study, essentially 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. We reproduce these two primary study results in our between-subjects replication exercise with an additional 200 UC Berkeley students. Our third finding is that 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 present bias measured in the first block of the experiment is predictive of this (later) commitment choice. A corresponding investigation on the extent of sophistication and commitment demand indicates that subjects potentially forecast 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 aware of their dynamic inconsistency as they demand binding commitment.

The paper proceeds as follows: Section 2 provides details for our longitudinal experimental design. Section 3 presents results and section 4 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 those 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. The between-subjects design is discussed primarily in section 3.5 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. Random Greek letters appear, 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.[5] 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.

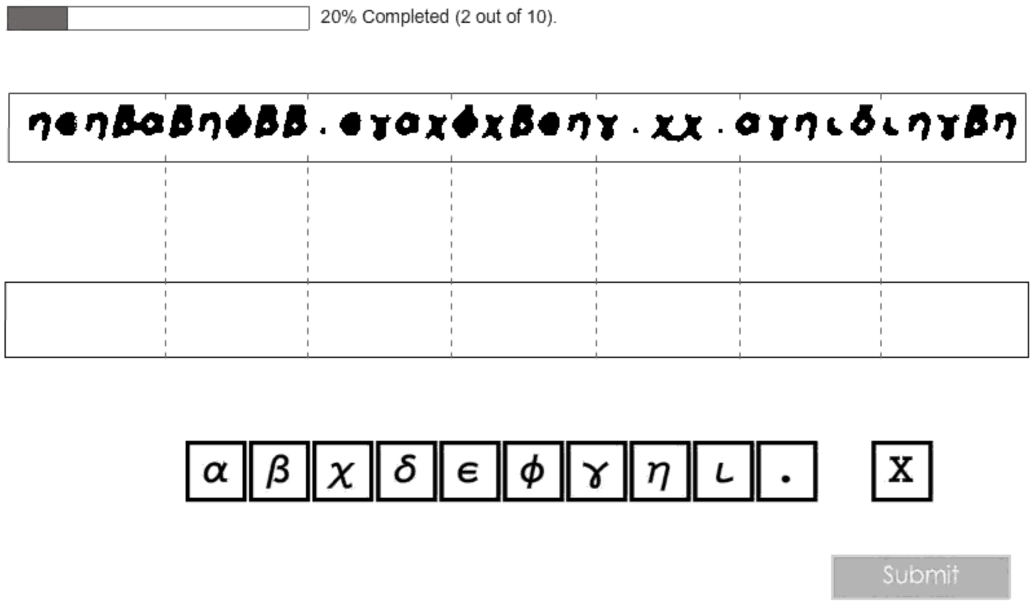
2.2 Experimental Timeline

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

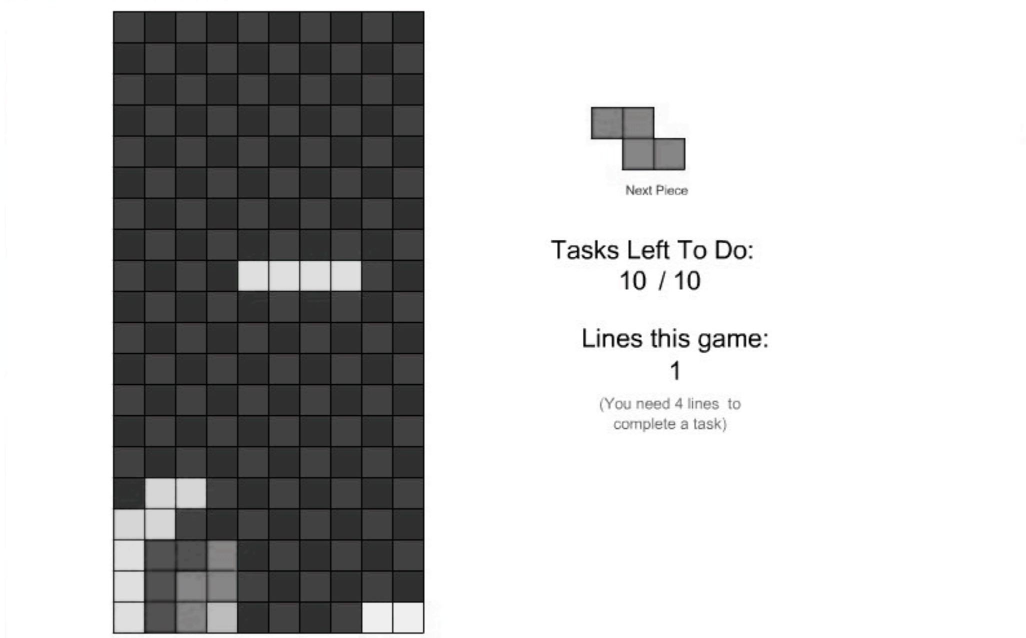
[5] Our measure of accuracy is the Levenshtein Distance. The Levenshtein Distance is commonly used in computer science to measure the distance between two strings and is defined as the minimum number of edits needed to transform one string into the other. Allowable edits are insertion, deletion or change 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



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 environment. Using slider bars, subjects allocate tasks to two dates, one earlier and one later, under different interest rates.[6] Figure 2 provides a sample allocation screen. To motivate the intertemporal tradeoffs faced by subjects, decisions are described as having different 'task rates.' 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 in Week 2.[7]

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.[8]

2.3.2 Minimum Work

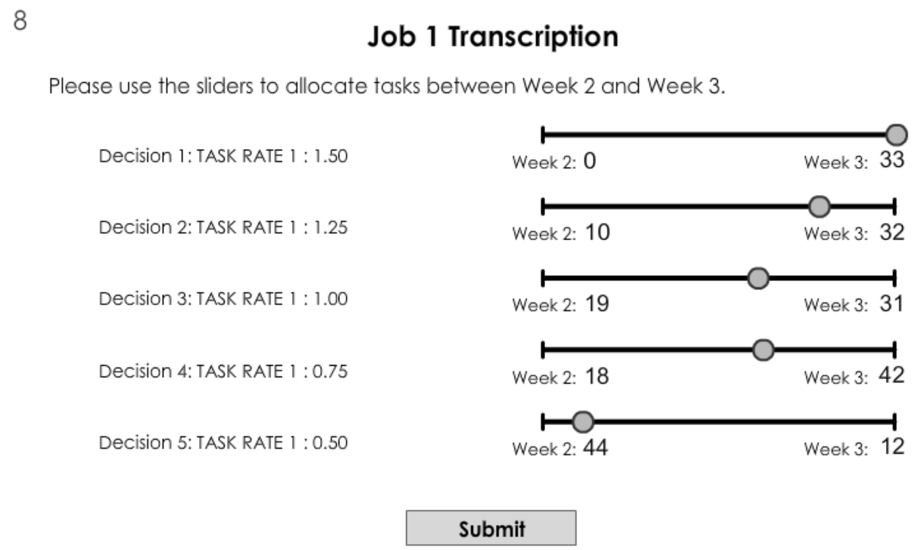
In each week, subjects are required to complete 10 tasks of each Job prior to making allocation decisions or completing allocated tasks. The objective of these required tasks, which we call

[6]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.

[7]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 3.5 for detail.

[8]We use R for present value budget constraints of the form $e_t + R \cdot e_{t+k} = m$, and P for future value budget constraints of the form $P \cdot e_t + e_{t+k} = m$.

Figure 2: Convex Allocation Environment



"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.[9] 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.

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

[9]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).

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.[10] 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.

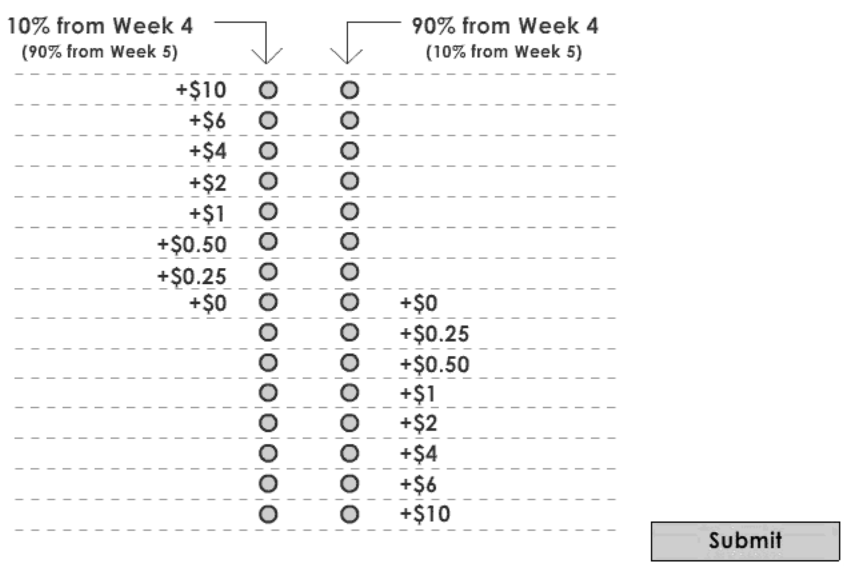
In order to operationalize our elicitation of commitment demand, subjects are asked to make 15 multiple price list decisions between two options. In the first option, the allocation-that-counts will come from Week 4 with probability 0.1. In the second option, the allocation-that-counts will come from 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.[11] Figure 3 provides the implemented price list. One of the 15 commitment

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

[11] 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.

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.[12] 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.[13] This isolation is encouraged as subjects are told to treat each decision as if it was the one going to be implemented (See Appendix F.4 for detail).

Figure 3: Commitment Demand Elicitation



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

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

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

likely to demand commitment? Third, a correlation between time inconsistency and commitment validates the interpretation of present bias over other explanations for time inconsistent 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 such surprises should favor flexibility to accommodate her noisy schedule. In contrast, 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.[14] 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.[15] One more subject completed initial allocations in Week 1, but due to computer error did not have their choices recorded. 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. In our sample, nine subjects have this issue for one or more weeks of the study.[16] For the analysis, we focus on

[14]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.

[15]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 qualitatively alter the analysis or conclusions.

[16]Appendix Tables A5 and A6 provide estimates for each individual based on their Block 1 data. The 9 individuals without variation in their responses in one or more weeks are noted. Extreme estimates are obtained for individuals without variation in experimental response in one of the weeks of Block 1.

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 A9, 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 allocate monetary payments across the two dates t and $t+k$, c_t and c_{t+k}, subject to the intertemporal 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).[17] 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.

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

Appendix F provides the full experimental instructions.

While the monetary discounting experiment replicates the design of Andreoni and Sprenger (2012a) to a large extent, there are two important differences. First, Andreoni and Sprenger (2012a) implement choices with payment by check. Our design implements payment by cash with potentially lower transaction costs. Second, Andreoni and Sprenger (2012a) implement choices with present payment received only by 5:00 p.m. in a subject's residence mailbox. If these payments are not construed as the present, one would expect no present bias. Here, we provide payment immediately in the lab.

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 Results

The results are presented in five 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. Fourth we present results related to commitment demand, documenting correlations with previously measured present bias and analyzing the value of commitment. Lastly, a fifth subsection is dedicated to a between-subjects replication exercise of the results concerning differences in discounting when comparing choices over monetary rewards to effort chocies.

3.1 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 from $P \cdot c_t + c_{t+k} = 20$ is reported for the 75 subjects from the primary sample for whom we have all monetary data. The left panel shows three data series 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.

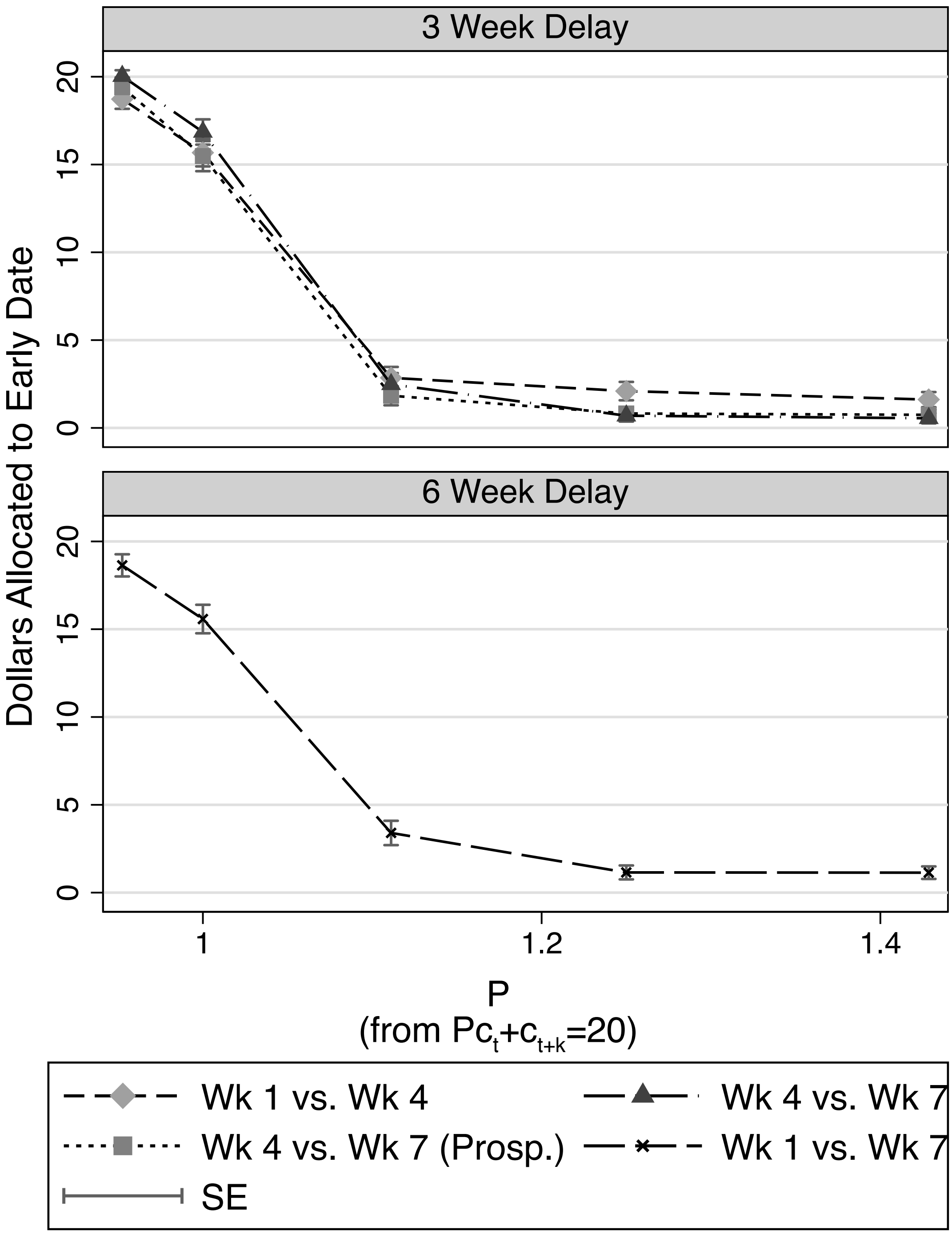
We highlight two features of Figure 4. First, note that as P from $P \cdot c_t + c_{t+k} = 20$ increases, the average allocation to the sooner payment decreases, following the law of demand. Indeed, at the individual level 98% of choices are monotonically decreasing in P, and only 1 subject exhibits more than 5 non-monotonicities in demand in their monetary choices.[18] 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.[19] 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 payments, this comparison provides a secondary measure of present

[18] 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.

[19] Though the six-week delay data are used in estimation, our non-parametric tests only identify present bias from choices over three-week delays. Without parametric assumptions for utility our data do not lend themselves naturally to the method of identifying present bias where short horizon choices are compared to long horizon choices to examine whether discount factors nest exponentially (see, for example Kirby, Petry and Bickel, 1999; Giordano, Bickel, Loewenstein, Jacobs, Marsch and Badger, 2002).

Figure 4: Monetary Discounting Behavior



bias. In this comparison, controlling for P, subjects allocate on average \$0.47 ($s.e = 0.32$) more to the sooner payment when it is in the present, $F(1, 74) = 2.08$, ($p = 0.15$).[20] Table 2, Panel A provides a corresponding tabulation of behavior, presenting the budget share allocated to the sooner payment date and the proportion of choices that can be classified as present-biased. Budget shares for the sooner payment are calculated as $(P \cdot c_t)/m$ for each allocation. Across all values of P subjects allocate around 38% ($s.e. = 1.73$) of their experimental budget to the sooner payment date when the sooner date is in the future ($t \neq 0$) and around 41% (1.34) to the sooner payment date when the sooner date is in the present ($t = 0$), $F(1, 74) = 3.50$, ($p = 0.07$). Further, across all values of P, seventy-eight percent of choices are dynamically consistent, 13% are present-biased, and 9% are future-biased.

We find limited non-parametric support for the existence of a present bias over monetary payments. To provide corresponding estimates of present bias we follow the parametric assumptions of Andreoni and Sprenger (2012a) and assume quasi-hyperbolic (Laibson, 1997; O'Donoghue and Rabin, 2001) power utility with Stone-Geary background parameters. Hence, the quasi-hyperbolic discounted utility from experimental payments at two dates, c_t, received at time t, and c_{t+k}, received at time $t + k$, is

$$U(c_t, c_{t+k}) = (c_t + \omega)^\alpha + \beta^{\mathbf{1}_{t=0}} \delta^k (c_{t+k} + \omega)^\alpha. \tag{3}$$

The variable $\mathbf{1}_{t=0}$ is an indicator for whether or not the sooner payment date, t, is the present. 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. The utility function is assumed to be concave, $\alpha < 1$, such that first order conditions provide meaningful optima. Here, ω is a Stone-Geary background parameter that we take to be the \$5 minimum payment

[20] 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.

Table 2: Aggregate Behavior By Interest Rate

Panel A: Monetary Choices						
P	$t \neq 0$ Budget Share (1)	$t = 0$ Budget Share (2)	t-test (p-value) (3)	Proportion Present-Biased (4)	Proportion Dynamically Consistent (5)	Proportion Future-Biased (6)
0.952	0.924	0.923	0.07	0.073	0.813	0.113
	(0.228)	(0.189)	(p=0.94)			
1	0.774	0.813	1.32	0.200	0.660	0.140
	(0.368)	(0.323)	(p=0.19)			
1.11	0.102	0.148	1.86	0.180	0.733	0.087
	(0.259)	(0.300)	(p=0.06)			
1.25	0.051	0.087	1.97	0.113	0.853	0.033
	(0.177)	(0.239)	(p=0.05)			
1.429	0.053	0.077	1.40	0.100	0.847	0.053
	(0.182)	(0.228)	(p=0.16)			
Overall	0.381	0.410	1.87	0.133	0.781	0.085
	(0.461)	(0.458)	(p=0.07)			
Panel B: Effort Choices						
R	Initial Budget Share (1)	Subsequent Budget Share (2)	t-test (p-value) (3)	Proportion Present-Biased (4)	Proportion Dynamically Consistent (5)	Proportion Future-Biased (6)
0.5	0.787	0.761	1.76	0.294	0.444	0.263
	(0.180)	(0.219)	(p=0.08)			
0.75	0.717	0.690	1.70	0.356	0.363	0.281
	(0.206)	(0.245)	(p=0.09)			
1	0.541	0.489	3.65	0.237	0.656	0.106
	(0.134)	(0.183)	(p<0.01)			
1.25	0.324	0.250	4.12	0.388	0.444	0.169
	(0.239)	(0.222)	(p<0.01)			
1.5	0.289	0.222	3.67	0.369	0.425	0.206
	(0.242)	(0.226)	(p<0.01)			
Overall	0.532	0.482	3.86	0.329	0.466	0.205
	(0.286)	(0.311)	(p<0.01)			

Notes: Panel A tabulates $t \neq 0$ and $t = 0$ budget shares for sooner payments for each P in money. Each row calculates from 75 $t \neq 0$ allocations (one at each interest rate in the Week 4 vs. Week 7 prospective choices) and 150 $t = 0$ allocations (one at each interest rate in the Week 4 vs. Week 7 actual and Week 1 vs. Week 4) choices. Paired t-tests with 149 degrees of freedom presented. Panel B tabulates initial and subsequent budget shares for sooner tasks for each R in effort. Each row calculates from 160 initial allocations (one each for tetris and greek at each task rate) and 160 subsequent allocations. Paired t-tests with 159 degrees of freedom presented. Overall tests in both panels come from regression of budget share on allocation timing with standard errors clustered on individual level. Test statistic is t-statistic testing the null hypothesis of no effect of allocation timing, which controls for multiple comparisons.

of the monetary experiment.[21] Maximizing (3) subject to the intertemporal budget constraint

[21] Andreoni and Sprenger (2012a) provide detailed discussion of the use of such background parameters and

(2) yields an intertemporal Euler equation, which can be rearranged to obtain

$$log(\frac{c_t + \omega}{c_{t+k} + \omega}) = \frac{log(\beta)}{\alpha - 1} \cdot (\mathbf{1}_{t=0}) + \frac{log(\delta)}{\alpha - 1} \cdot k + (\frac{1}{\alpha - 1}) \cdot log(P). \tag{4}$$

Assuming an additive error, this functional form can be estimated at the aggregate or individual level.[22] One important issue to consider in estimation 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 (for discussion, see Wooldridge, 2002; Andreoni and Sprenger, 2012a). Discounting and utility function parameters can be recovered via non-linear combinations of regression coefficients with standard errors estimated via the delta method. Appendix A provides a detailed discussion of identification and estimation of discounting parameters for both monetary and effort choices. [23]

In Table 3, columns (1) and (2) we implement two-limit Tobit regressions 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 our non-parametric evidence. Across specifications we identify weekly discount factors of around 0.99. The 95% confidence interval in column (1) for the weekly discount factor implies annual discount rates

---

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.

[22] An additive error yields the regression equation

$$log(\frac{c_t + \omega}{c_{t+k} + \omega}) = \frac{log(\beta)}{\alpha - 1} \cdot (\mathbf{1}_{t=0}) + \frac{log(\delta)}{\alpha - 1} \cdot k + (\frac{1}{\alpha - 1}) \cdot log(P) + \epsilon.$$

The stochastic error term, ϵ, is necessary to rationalize any discrepancies between our theoretical development and our experimental data. One simple foundation for such an error structure would be to assume that individuals exhibit random perturbations to their log allocation ratios, $log(\frac{c_t+\omega}{c_{t+k}+\omega})$. A more complete formulation might follow macroeconomic exercises such as Shapiro (1984), Zeldes (1989), and Lawrance (1991). With a time series of consumption, one assumes rational expectations such that Euler equations are satisfied up to a mean zero random error, uncorrelated with any information available to the decisionmaker. Assuming constant relative risk aversion, as we do, this forecast error provides the structure for estimating utility function curvature and recovering discounting parameters in a way very similar to our exercise.

[23] The notation of Appendix A is slightly altered to discuss allocation timing and make links to partial sophistication and the value of commitment for effort choices.

Table 3: 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)
Weekly Discount Factor: $(\delta)^7$	0.988	0.980	0.993	1.007	0.999
	(0.003)	(0.003)	(0.027)	(0.029)	(0.025)
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
$H_0 : \beta = 1$	$\chi^2(1) = 8.77$	$\chi^2(1) = 1.96$	$\chi^2(1) = 7.36$	$\chi^2(1) = 11.43$	$\chi^2(1) = 11.42$
	$(p < 0.01)$	$(p = 0.16)$	$(p < 0.01)$	$(p < 0.01)$	$(p < 0.01)$
$H_0 : \beta(Col.\ 1) = \beta(Col.\ 5)$	$\chi^2(1) = 6.37$				
	$(p = 0.01)$				
$H_0 : \beta(Col.\ 2) = \beta(Col.\ 5)$		$\chi^2(1) = 8.27$			
		$(p < 0.01)$			

Notes: Parameters identified from two-limit Tobit regressions of equations (4) and (6) 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). Chi-squared tests inlast three rows.

between 40% and 140%.[24] In column (1) of Table 3 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 delays, we find $\beta = 0.988$ (0.009) and are unable to reject the null hypothesis of dynamic consistency, $H_0 : \beta = 1$: $\chi^2(1) = 1.96$, $(p = 0.16)$. These estimates demonstrate limited present bias for money and hence confirm the non-parametric results.

In both specifications, we estimate α of around 0.975 indicating limited utility function

[24]In Appendix A, we discuss identification of all parameters and note that discount factors are identified from variation in delay length, k. Our ability to precisely identify aggregate discounting was not a focus of the experimental design and is compromised by limited variation in delay length. In monetary discounting experiments it is not unusual to find implied annual discount rates in excess of 100%.

curvature over monetary payments. Finding limited curvature over money is important in its own right, as linear preferences over monetary payments are 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, note that 86% of monetary allocations are corner solutions and 61% of subjects have zero interior allocations in twenty decisions.[25]

Our non-parametric and parametric results closely mirror the aggregate findings of Andreoni and Sprenger (2012a) and Gine et al. (2010).[26] 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.

To summarize, 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 (recall that we estimate α to be around 0.975). 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.

[25] A consequence of limited utility function curvature is that even a small degree of present bias can lead potentially to sizable changes in allocation behavior through time as individuals may switch from one corner solution to another. Hallmarks of this are seen in Table 2, which tabulates behavior across interest rates. Though a wide majority of observations are dynamically consistent, some significant changes in budget shares are seen at specific interest rates.

[26] In both of these prior exercises substantial heterogeneity in behavior is uncovered. In subsection 3.3 we conduct individual analyses, revealing similar findings.

3.2 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 E.5 we present results of present bias from both blocks of the experiment and document very similar findings.

In Figure 5, we show for each value of R from $e_t + R \cdot e_{t+k} = 50$, the amount of tasks allocated to the sooner work date, Week 2, which could range from 0 to 50.[27] We contrast initial allocations of effort made in Week 1 with subsequent allocations made in Week 2 for the 80 subjects of the primary sample. Standard error bars are clustered at the individual level.

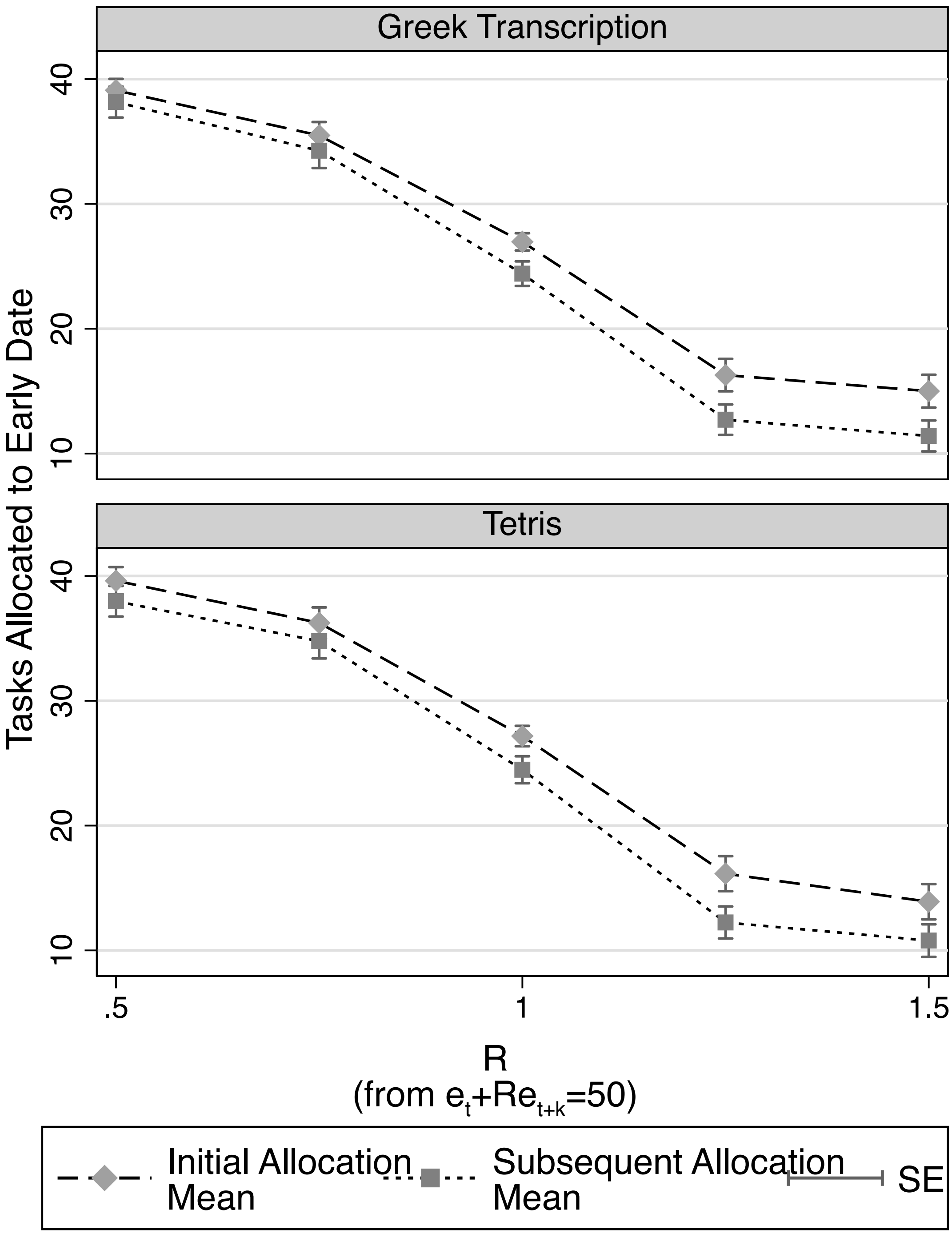
As with monetary discounting, subjects appear to have understood the central intertemporal tradeoffs of the experiment as both initial and subsequent allocations decrease as R is increased. At the individual level, 95% of choices are monotonically decreasing in R, and only 5 subjects exhibit more than 5 non-monotonicities in their effort choices.[28] This suggests that subjects as a whole understand the implied intertemporal tradeoffs and the decision environment.

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 $F(1, 79) =$

[27]The data are presented as a function of R from $e_t + R \cdot e_{t+k} = 50$, as opposed to relative price, to provide a standard downward sloping demand curve. Recall that $R \in \{0.5, 0.75, 1, 1.25, 1.5\}$. When R is low, sooner tasks are relatively cheap to complete, and when R is high, sooner tasks are relatively expensive to complete.

[28]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, $F(1, 79) = 8.34$ ($p < 0.01$).

Figure 5: Real Effort Discounting Behavior



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).[29] Table 2, Panel B provides a corresponding tabulation of behavior, presenting the budget share allocated to the sooner work date and the proportion of choices that can be classified as present-biased. Budget shares for the sooner work date are calculated as e_t/m for each allocation. Across all values of R, subjects initially allocate around 53% ($s.e. = 0.97$) of their experimental budget to the sooner work date and subsequently allocate around 48% (1.02) to the sooner work date, when that sooner work date is in the present, $F(1, 79) = 14.87$, ($p < 0.01$). Across all values of R, forty-seven percent of choices are dynamically consistent, 33% are present-biased, and 21% are future-biased.[30]

Motivated by our non-parametric analysis we proceed to estimate intertemporal parameters. Subjects allocate effort to an earlier date, e_t, and a later date, e_{t+k}. We again assume quasi-hyperbolic discounting and a stationary power cost function with Stone-Geary background parameters to write the discounted costs of effort as

$$(e_t + \omega)^\gamma + \beta^{\mathbf{1}_{t=0}} \delta^k (e_{t+k} + \omega)^\gamma. \tag{5}$$

Here $\gamma > 1$ represents the stationary parameter on the convex instantaneous cost of effort function. The Stone-Geary term, ω, could be interpreted as some background level of required work. For simplicity, we interpret ω as the required minimum work of the experiment and set $\omega = 10$ for our effort analysis. The variable $\mathbf{1}_{t=0}$ is an indicator for whether or not the sooner work date, t, is the present. As before, the parameter β captures the degree of present bias and the parameter δ captures long run discounting.

Maximizing (5) subject to (1) ($e_{t,t} + R \cdot e_{t+k,t} = 50$) yields an intertemporal Euler equation,

[29] 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 E.5 for detail.

[30] Appendix Table A3 provides identical analysis using both blocks of data and reports very similar results.

which can be rearranged to obtain

$$log(\frac{e_t+\omega}{e_{t+k}+\omega}) = \frac{log(\beta)}{\gamma-1}\cdot(\mathbf{1}_{t=0}) + \frac{log(\delta)}{\gamma-1}\cdot k - (\frac{1}{\gamma-1})\cdot log(R). \tag{6}$$

As before, we assume an additive error structure and estimate the linear regression implied by (6) using two-limit Tobit regression. The parameters of interest are again recovered from non-linear combinations of regression coefficients with standard errors calculated via the delta method. Appendix A provides detailed discussion of identification for such choices.[31]

Table 3 columns (3) through (5) present two-limit Tobit regressions with standard errors clustered on the individual level. 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.[32] 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 weekly discount factor of $\delta = 0.993$ is similar to our findings for monetary discounting.

In column (3) of Table 3 we estimate an aggregate $\beta = 0.900$ (0.037), and 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 modified 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.[33] 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

[31] The notation of Appendix A is slightly altered to discuss allocation timing and make links to partial sophistication and the value of commitment for effort choices.

[32] 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.

[33] 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 A10 provides estimates.

Week 2 work date are around eleven percent per week. We therefore confirm our non-parametric findings on effort choices.

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), $\chi^2(1) = 8.27,\ (p < 0.01)$. Subjects are significantly more present-biased over effort than over money.[34]

3.3 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 3.[35] These regressions assume no heterogeneity in cost or utility function curvature and recover individual parameter estimates of β_e, present bias for effort, and β_m, present bias for money, as non-linear combinations of regression coefficients. The methods for identifying individual discounting parameters are discussed in Appendix A.[36] Appendix Tables A5 and A6 provide individual estimates of β_e and β_m along with a summary of allocation behavior for both effort and money for each subject.[37]

[34]In Appendix E.5 we conduct identical analysis using both Blocks 1 and 2 and arrive at the same conclusions. See Appendix Table A11 for estimates.

[35]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 3, column (1).

[36]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. Nearly identical aggregate discounting estimates are generated when conducting ordinary least squares versions of Table 3. Curvature estimates, however, are sensitive to estimation techniques that do and do not recognize that the tangency conditions implied by (6) and (4) may be met with inequality at budget corners. Se Andreoni and Sprenger (2012a) for further discussion.

[37]Appendix Tables A5 and A6 include data from the 9 subjects excluded from the primary study sample for

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, $z = 2.31$, ($p = 0.02$).[38]

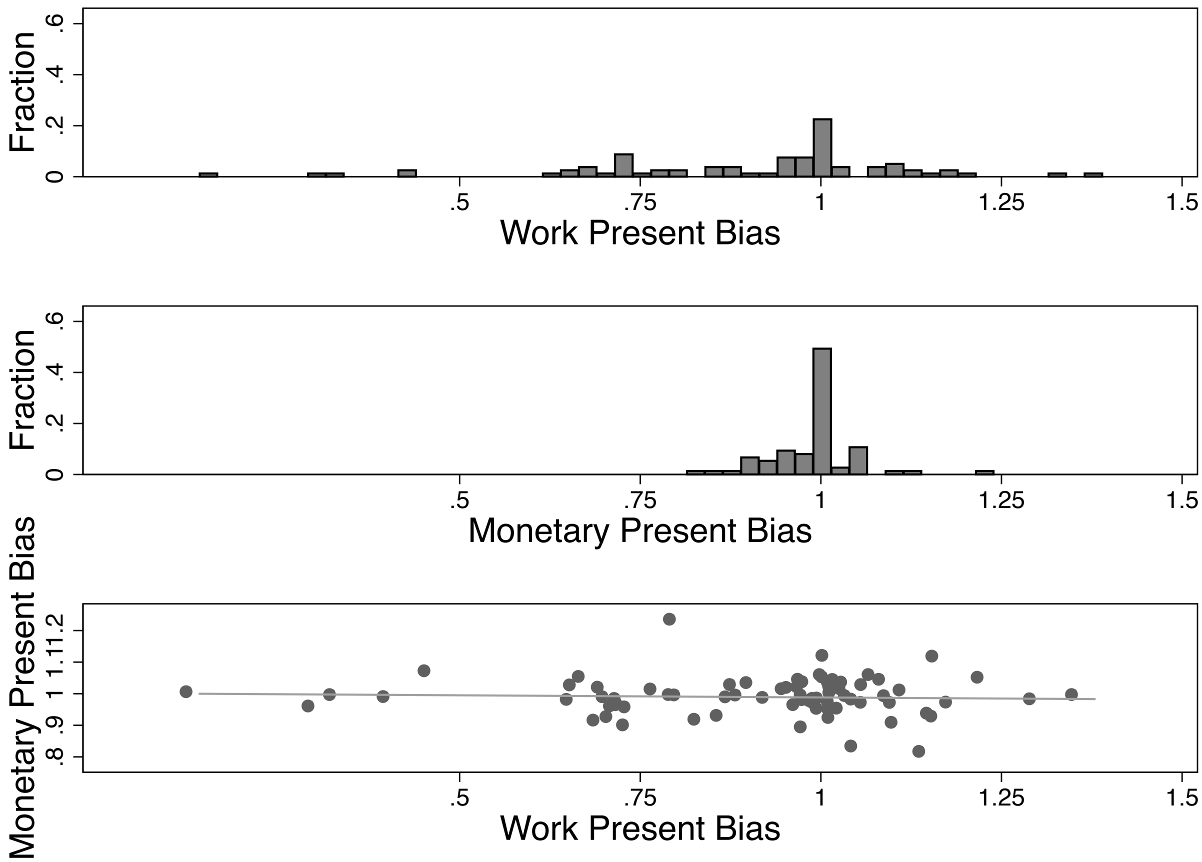
Two important questions with respect to our individual measures arise. First, how much do these measures correlate within individual? The answer to this question is important for understanding both the validity of studies relying on monetary measures and the potential consistency of preferences across domains. Significant correlations would suggest that there may be some important preference-related behavior uncovered in monetary discounting studies.[39] 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,

having no variation in experimental response in one or more weeks of the study. These subjects are noted along with an explanation of which weeks they provided no variation in response.

[38]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.

[39]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, $\rho = 0.11$, $(p = 0.33)$.[40]

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 difference measures for allocations. For effort choices we calculate the budget share of each allocation for Week 2 effort.

[40]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.

The difference in budget shares between subsequent allocation and initial allocation is what we term a 'budget share difference.'[41] As budget shares are valued between $[0,1]$, our difference measure takes values on the interval $[-1,1]$. Negative numbers indicate present-biased behavior and values of zero indicate dynamic consistency. Each subject has 10 such effort budget share difference measures in Block 1. The average budget share difference for effort is -0.049 (s.d. = 0.115) indicating that subjects allocate around 5% less of their work budget to the sooner work date when allocating in the present.[42] At the individual level, 49 of 80 subjects have an average budget share difference of less than zero, 13 have an average difference of exactly zero, and 18 have an average difference greater than zero, demonstrating a modal pattern of present bias.

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 difference measures. The average budget share difference for money is -0.029 (s.d. = 0.134).[43] At the individual level, 28 of 75 subjects have an average budget share difference of less than zero, 32 have an average difference of exactly zero, and 15 have an average difference greater than zero, demonstrating a modal pattern of dynamic consistency.

The non-parametric budget share difference measures are closely correlated with our parametric estimates at the individual level. The correlation between β_e and each individual's average budget share difference for effort is $\rho = 0.948$, $(p < 0.01)$. Of the 49 individuals with negative average budget share differences for effort, 47 have estimates of $\beta_e < 1$. Of the 18 individuals with positive average budget share differences for effort, all 18 have estimates of $\beta_e > 1$.

[41] 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 difference is $\frac{e_2'-e_2}{50}$.

[42] As noted previously, this average value deviates significantly from the dynamically consistent benchmark of 0, $F(1,79) = 14.87$, $(p < 0.01)$.

[43] As noted previously, this average value differs marginally significantly from the dynamically consistent benchmark of 0, $F(1,74) = 3.50$, $(p = 0.07)$.

Of the 13 individuals with zero average budget share differences for effort, 11 have $\beta_e = 1$ and 2 have $\beta_e = 1.003$. The correlation between β_m and each individual's average Budget Share Difference for money is $\rho = 0.997$, ($p < 0.01$). Of the 28 individuals with negative average budget share differences for money, all 28 have estimates of $\beta_m < 1$. Of the 15 individuals with positive average budget share differences for money, all 15 have estimates of $\beta_m > 1$. Of the 32 individuals with zero average budget share differences for money, all 32 have $\beta_m = 1$.[44] This apparent internal validity 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. However, many alternative explanations exist for rationalizing these data patterns. Chief among these alternatives are the existence of unanticipated shocks to the cost of performing tasks (either in general or specific to tasks in Week 2), resolving uncertainty between allocation times, and subject exhaustion or error. These alternative explanations are considered in detail in Appendix C. Importantly, we show in Appendix C that under none of these alternatives would we expect a clear link between the behavioral pattern of reallocating fewer tasks to the present and commitment demand. 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 to measured present bias.

3.4 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

[44] Appendix Tables A5 and A6 provide all the corresponding estimates and average budget share data.

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 subsection 3.4.1 we explore the value of commitment and choices when X or Y are not zero. In the simple choice where neither commitment nor flexibility were costly, 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.

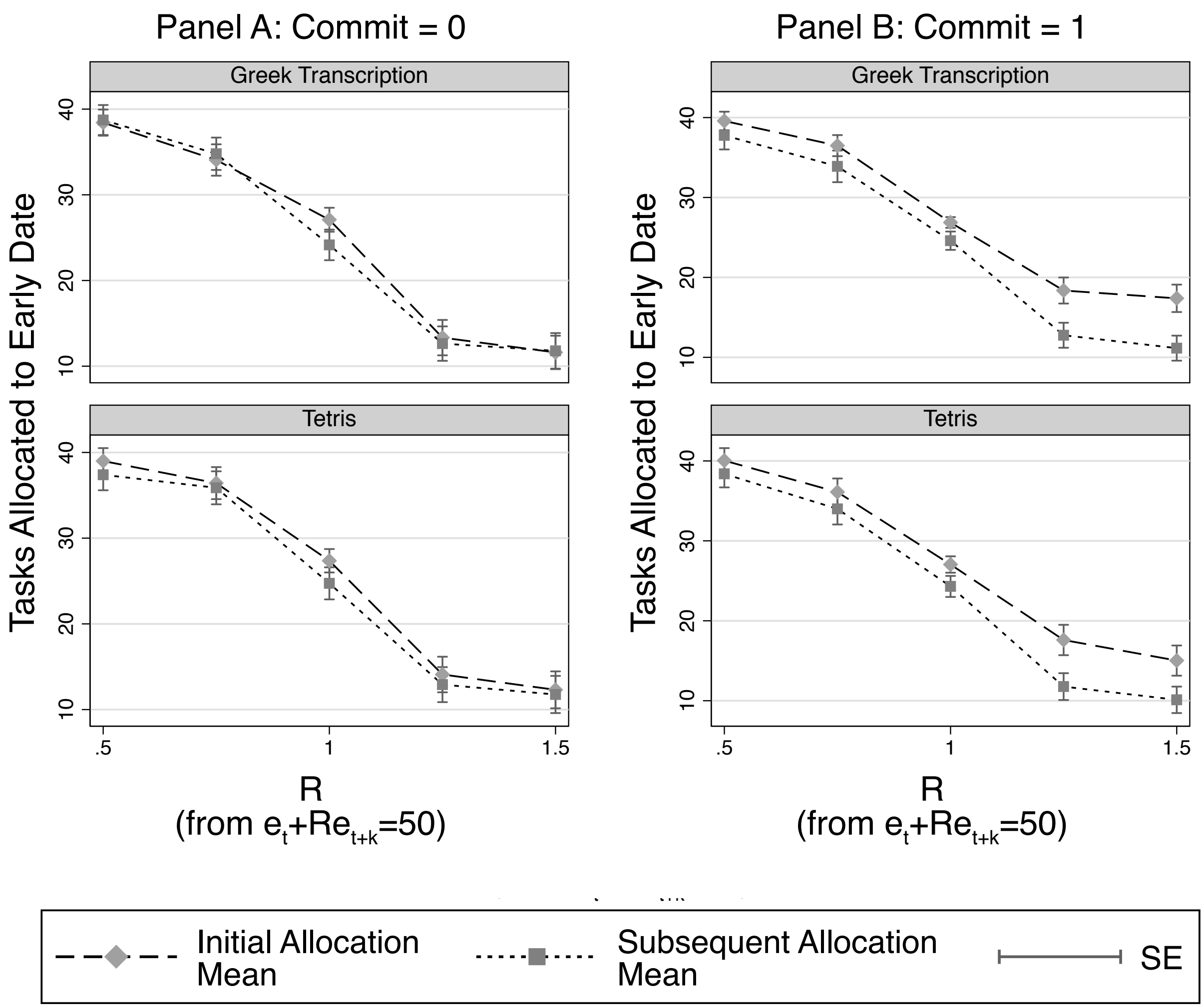
Figure 7 presents Block 1 task allocation behavior separated by commitment choice in Block 2. Immediately apparent from Figure 7 is that experimental behavior separates along commitment choice. Subjects who choose commitment in Week 4 made substantially present-biased task allocations in Week 2 given 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 who demand commitment in Week 4 altered their allocations by significantly more tasks than subjects who did not demand commitment, $F(1,79) = 5.84,\ (p = 0.02)$.[46]

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

[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.

[46] When including the 9 subjects with insufficient variation, this relationship between commitment and present-biased reallocations is no longer significant. Committers reallocate 0.90 (clustered s.e. = 1.32) fewer tasks to the sooner work date when the sooner work date is the present compared to non-committers, $F(1,88) = 0.46,\ (p = 0.49)$. We believe this is due to the fact that the nine subjects with insufficient variation lie at the extremes of changes in allocations in Block 1. Two of the nine would lie below the 5th percentile in budget share differences (leading to β_e estimates of 0.24 and 0.25) and one would lie above the 95th percentile (leading to a β_e estimate of 2.63). Removing these three extreme subjects, we find that committing subjects reallocate 2.19 (1.12) fewer tasks to the sooner work date when it is the present compared to non-committers, $F(1,88) = 3.86,\ (p = 0.05)$.

Figure 7: Commitment Choice and Allocation Behavior



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, $\chi^2(1) = 4.85$, $(p = 0.03)$.[47]

[47]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 A12 provides analysis including the data from both blocks. It is worth noting that the estimates of weekly discount factors, δ also differ across committing and non-committing subjects. This difference is identified from differences in initial allocations. Non-committing subjects have an average initial budget share for sooner tasks of 50.7% (clustered s.e. = 1.6) and an average subsequent budget share of 49.0% (1.7), while committing subjects have an average initial budget share of 54.9% (1.3) and an average subsequent share of 47.7% (1.4). Committing subjects' behavior is consistent with $\delta > 1$. However, we hesitate to draw any firm conclusions from this observation

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, ($p = 0.26$).

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)
Weekly Discount Factor: $(\delta)^7$	0.978 (0.003)	0.981 (0.005)	0.917 (0.032)	1.065 (0.039)
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 two-limit Tobit regressions of equations (4) and (6) 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 $H_0 : \beta(Col.\ 3) = \beta(Col.\ 4)$.

as our experiment provides no variation in delay lengths to help identify δ. As discussed in Appendix A, δ is identified from the constant one week delay between work dates. Hence, any level differences across subjects are revealed as differences in estimated δ parameters.

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 that correlation is far from perfect. For example, the raw correlation between β_e and commitment choice is $\rho = 0.225$, $(p = 0.04)$, implying an R-squared value of around 5%. Substantial variance in the choice of commitment remains unexplained. There are several potential reasons for this lack of explanatory power. A natural first possibility is substantial naivete. Though our results suggest at least partial sophistication, on average, many subjects may be naive with respect to their dynamic inconsistency. Further, among partially sophisticated individuals, there may be limited correlation between behavior and beliefs such that individuals with both high and low values of β_e may share similar beliefs as to their future behavior. Third, there may be uncertainty in the work environment uncontrolled by the researcher. Even sophisticated present-biased individuals may wish to remain flexible. In subsection 3.4.1 and Appendix D we discuss uncertainty and the benefits of flexibility in detail, noting that the value of commitment is likely influenced by the unmodeled benefits of flexibility. Fourth, the allocation decisions may be subject to substantial noise, leading at least partially to a misestimation of preferences and a misclassification of subjects. Each of these forces may be at play to certain degree, reducing our ability to tightly measure present bias and the extent of sophistication. However, our finding of a significant present bias and a correlation between present bias and commitment demand points to at least partial sophistication for some subjects.

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 instantaneously desire?[48] Given the nature of our commitment device, commitment will bind

[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

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 differences for Block 1. Non-committers have a mean budget share difference of -0.018 (clustered s.e. $=$ 0.009) allocating about 2 percentage points less of each budget to Week 2 when deciding in the present. In contrast, committers have a mean budget share difference 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 difference of 0.011 (0.017) while committers have a mean difference 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,

predicted link between sophisticated present bias and commitment demand.

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

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. 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.

3.4.1 The Value of Commitment

A natural question is how much should subjects be willing to pay for commitment. In Appendix A 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 price of commitment 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 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 and non-committers, the median valuation is \$0.125 and \$-0.125, respectively.

What do these monetary valuations imply for the extent of V and correspondingly for the extent of sophistication? In Appendix A, we theoretically investigate the valuation of commitment through the lens of the partially sophisticated quasi-hyperbolic model of O'Donoghue and Rabin (2001). We recover the valuation of commitment, V, for stationary cost functions. This analysis shows that the value of commitment is linked to the extent of sophistication, which is governed by sophistication parameter $\widehat{\beta}$, reflecting an individual's assessment of their future present bias. If $\widehat{\beta} = 1$, an individual is perfectly naive, and if $\widehat{\beta} = \beta$, an individual is perfectly sophisticated. Values of $\widehat{\beta} \in \{\beta, 1\}$ correspond to partial sophistication. That present bias is predictive of commitment demand at price zero indicates at least partial sophistication on average, $\widehat{\beta} < 1$.

The level of V can be calculated directly for the fully sophisticated benchmark of $\widehat{\beta} = \beta$, which implies a perfect forecast for present-biased behavior. 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.

To relate the value of roughly two tasks to mooney, note that on average, using minimum work completion rate, subjects 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

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

[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 D analyzes the value of commitment demand at a wide range of potential per task valuations to provide sensitivity analysis.

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.[52]

We are hesitant to draw strong conclusions beyond the plausibility 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.

3.5 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 interpretation is that models of dynamic inconsistency are validated when tested in their relevant domain (consumption) and that choices over fungible monetary payments cannot easily speak to such models' predictions.

However, in our 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.[53] 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 money from a combination

[52] If individuals are fully sophisticated, monetary valuations for commitment should be close to 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 D 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.

[53] That is, the constraint for effort was of a present value form, $e_t + Re_{t+k} = 50$, while the constraint for money was of a future value form, $Pc_t + c_{t+k} = 20$.

of static and dynamic choices.[54] 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.[55]

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

$$Pa_2 + a_3 = m.$$

[54]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.

[55]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.

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 $P \in \{0.66, 0.8, 0.91, 0.95, 1, 1.05, 1.11, 1.25, 1.54\}$, 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 and minimum work of 10 tasks 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 G 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, we can directly compare present bias across conditions. Figure 8 plots the amount of money in Panel A (out of \$20) or the number of tasks in Panel B (out of 60) and allocated to Week 3 for each level of P. 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 P, while Week 3 money is increasing in P. Note as well that due to the form of the budget, it is the constant-value Week 3 units that are graphed.[56]

Figure 8 closely reproduces our prior within-subject findings. For money mean behavior appears almost perfectly dynamically consistent. Controlling for P, subjects allocate \$0.14 (clustered s.e. = 0.12) less to Week 3 in Week 2 relative to Week 1, $F(1, 98) = 1.37$, $p = 0.25$. In contrast, at each value of P, individuals appear present-biased for effort, allocating more effort to the later date when the sooner date is the present. Controlling for P, subjects allocate 2.14 (clustered s.e. = 1.10) more tasks to Week 3 in Week 2 relative to Week 1, $F(1, 94) = 3.82$, $p = 0.05$. Appendix Table A4 provides a corresponding tabulation of behavior, presenting budget shares and the proportion of choices that can be classified as present-biased.[57]

Non-parametric replication in hand, we now turn to estimation of aggregate utility parameters. In Table 5, we replicate the estimation exercise of Table 3 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 3, 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 5, column (1) and Table 3, column (2). For effort present bias in column (2), we estimate $\beta = 0.892$ (0.056), which compares favorably to Table 3, 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 5, column (2) and

[56] This is in contrast to the prior effort figures where earlier tasks had constant value and were graphed and the prior money figures where earlier money was also graphed for ease of comparison.

[57] For consistency with Table 2 and Appendix Table A3, Appendix Table A4 tabulates budget shares for the *sooner* date, calculated as $(Pa_2)/m$ for each allocation. For money, subjects initially allocate around 51.4% (0.7) of their experimental budget to the sooner payment and subsequently allocate around 51.9% (0.6) to the sooner payment, $F(1, 98) = 0.85$, $(p = 0.36)$. Eighty-three percent of individual choices are dynamically consistent, 10% are present-biased, and 7% are future-biased. For effort, subjects initially allocate around 52.4% (clustered s.e. = 1.1) of their experimental budget to the sooner work date and subsequently allocate around 48.8% (1.7) to the sooner work date, $F(1, 94) = 3.82$, $(p = 0.05)$. Twenty-five percent of individual choices are dynamically consistent, 43% are present-biased, and 32% are future-biased.

Figure 8: Between Subjects Replication Exercise

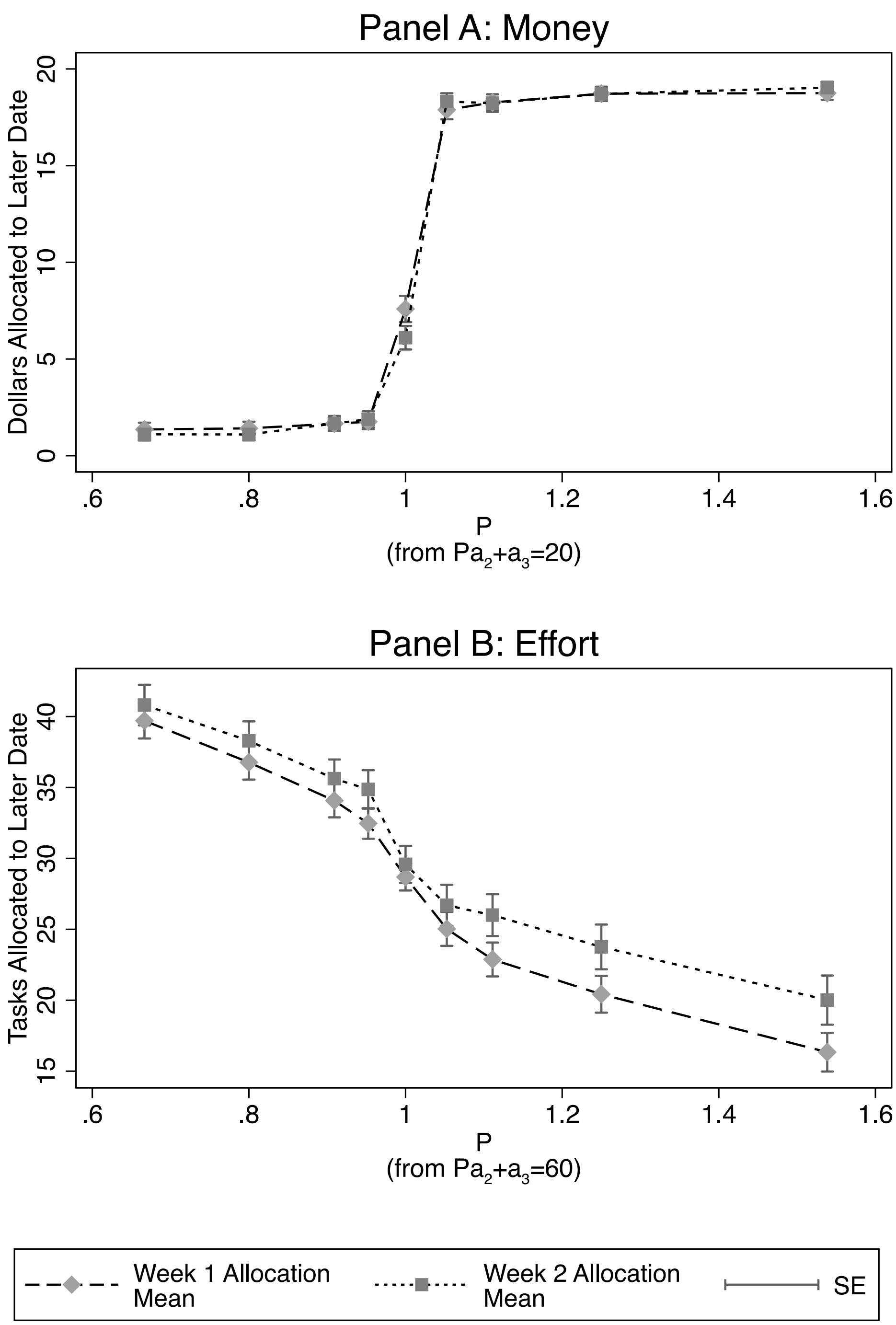


Table 3, column (3). The analysis again allows us to compare present bias across effort and money, and again we reject the null hypothesis that the β identified for money is equal to that identified for effort, $\chi^2(1) = 3.50,\ (p = 0.06)$.[58]

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 of the point estimates from our prior work.

Table 5: Replication Exercise Parameter Estimates

	Monetary Discounting	Effort Discounting Greek
	(1)	(2))
Present Bias Parameter: β	0.997 (0.005)	0.892 (0.056)
Weekly Discount Factor: $(\delta)^7$	0.998 (0.001)	1.009 (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
$H_0 : \beta = 1$	$\chi^2(1) = 0.50$ $(p = 0.48)$	$\chi^2(1) = 3.73$ $(p = 0.05)$
$H_0 : \beta(Col.\ 1) = \beta(Col.\ 2)$	$\chi^2(1) = 3.50$ $(p = 0.06)$	

Notes: Parameters identified from two-limit Tobit regressions of equations (4) and (6) 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. Chi-squared tests used in last two rows.

[58] Appendix Tables A7 and A8 provide individual estimates of β_e and β_m along with a summary of allocation behavior for these subjects. Subjects with no variation in experimental response in a given week are also noted. 16 of 194 non-attriting subjects have no variation in experimental response in one or more weeks and 14 of these subjects were in the effort condition. Importantly, the results of Table 5 are maintained if we eliminate such subjects with no variation in one or more weeks. See Appendix Table A13 for detail.

4 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 potentially 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).[59] 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 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. 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.

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

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.

Bauer, Michal, Julie Chytilova, and Jonathan Morduch, "Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India," *American Economic Review*, 2012, *102* (2), 1118–1139.

Benjamin, Daniel, James Choi, and Joshua Strickland, "Social Identity and Preferences," *American Economic Review*, 2010, *100* (4), 1913–1928.

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.

_ , _ , _ , and _ , "Are Risk Aversion and Impatience Related to Cognitive Ability," *American Economic Review*, 2010, *100* (3), 256–271.

Dupas, Pascaline and Jonathan Robinson, "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *American Economic Review*, 2013, *103* (4), 1138–1171.

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.

Giordano, Louis A., Warren K. Bickel, George Loewenstein, Eric A. Jacobs, Lisa Marsch, and Gary J. Badger, "Mild opioid deprivation increases the degree that opioid dependent outpatients discount delayed heroin and money," *Psychopharmacology*, 2002, *163*, 174–182.

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.

Kirby, Kris N., Nancy M. Petry, and Warren K. Bickel, "Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls," *Journal of Experimental Psychology: General*, 1999, *128*, 78–87.

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.

Lawrance, Emily C., "Poverty and the Rate of Time Preference: Evidence from Panel Data," *Journal of Political Economy*, 1991, *99* (1), 54–77.

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.

Shapiro, Matthew D., "The Permanent Income Hypothesis and the Real Interest Rate: Some Evidence from Panel Data," *Economics Letters*, 1984, *14* (1), 93–100.

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.

Sutter, Matthias, Martin G. Kocher, Daniela Glatzle-Ruetzler, and Stefan T. Trautmann, "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior," *American Economic Review*, 2013, *103* (1), 510–531.

Tanaka, Tomomi, Colin Camerer, and Quang Nguyen, "Risk and time preferences: Experimental and household data from Vietnam," *American Economic Review*, 2010, *100* (1), 557–571.

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

Voors, Maarten, Eleonora Nillesen, Philip Verwimp, Erwin Bulte, Robert Lensink, and Daan van Soest, "Violent Conflict and Behavior: A Field Experiment in Burundi," *American Economic Review*, 2012, *102* (2), 941–964.

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

Zeldes, Stephen B., "Consumption and Liquidity Constraints: An Empirical Investigation," *The Journal of Political Economy*, 1989, *97* (2), 305–346.