How soon is now? Evidence of present bias from convex time budget experiments

Uttara Balakrishnan, Johannes Haushofer & Pamela Jakiela

Experimental Economics
A Journal of the Economic Science Association

ISSN 1386-4157
Volume 23
Number 2

Exp Econ (2020) 23:294-321
DOI 10.1007/s10683-019-09617-y

Experimental Economics (2020) 23:294–321
https://doi.org/10.1007/s10683-019-09617-y

ORIGINAL PAPER

How soon is now? Evidence of present bias from convex time budget experiments

Uttara Balakrishnan[1] · **Johannes Haushofer**[2,3] · **Pamela Jakiela**[4]

Received: 24 September 2018 / Revised: 3 June 2019 / Accepted: 8 June 2019 / Published online: 29 April 2020

Abstract

We conduct a laboratory experiment in Kenya in which we elicit time and risk preference parameters from 494 participants, using convex time budgets and tightly controlling for transaction costs. Using the Kenyan mobile money system M-Pesa to make real-time transfers to subjects' phones , we vary whether same-day payments are made immediately after the experimental session or at the close of the business day. We find strong evidence of present bias, with estimates of the present bias parameter ranging from 0.902 to 0.924—but only when same-day payments are made immediately after the experiment.

Keywords Discount rate · Present bias · Experiment · Mobile money

JEL Classification C91 · D90 · O12

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10683-019-09617-y) contains supplementary material, which is available to authorized users.

✉ Pamela Jakiela
pjakiela@cgdev.org

Uttara Balakrishnan
uttara.balakrishnan@gmail.com

Johannes Haushofer
haushofer@princeton.edu

1 IMPAQ International, Washington, DC, USA

2 Princeton University, Princeton, USA

3 Busara Center for Behavioral Economics, Nairobi, Kenya

4 Center for Global Development, Washington, DC, USA

1 Introduction

How people trade off immediate and delayed consumption is a question of fundamental importance in economics (von Bohm-Bawerk 1890; Fisher 1930). The canonical economic model of time preferences is the discounted utility model, first proposed by Samuelson (1937); in it, all future payments are discounted by a constant factor each period, leading to exponential discounting.[1] In the second half of the twentieth century, the discounted utility model was called into question: empirically observed discounting behavior, both in animals and humans, did not correspond to the predictions of exponential discounting; in particular, short-term discount rates were found to be higher than long-term discount rates (Ainslie 1975; Thaler 1981). These findings led to the development of alternative models of intertemporal tradeoffs in which agents are present-biased in the sense that they overweight immediate payments relative to those that occur in the future.[2] In economics, the most widely used example is the quasi-hyperbolic model, first proposed by Phelps and Pollak (1968) and adapted to the case of time preferences by Laibson (1997) and O'Donoghue and Rabin (1999).[3] The quasi-hyperbolic model has since been used to explain empirical phenomena ranging from retirement saving (Laibson et al. 1998) to gym attendance (DellaVigna and Malmendier 2006; Acland and Levy 2015). Present bias is important in a range of policy settings because it predicts preference reversals: agents who exhibit present bias will make consumption and savings plans that they fail to carry out; more generally, present-biased agents tend to invest less than they intend to in goods that yield long-run benefits (e.g. education and exercise), and to consume more than they intend to when goods are associated with future costs (e.g. unhealthy foods).

In recent years, just as present bias has begun receiving widespread attention from policymakers (cf. World Bank 2015), some scholars have come to question the experimental evidence documenting violations of the discounted utility model. On the one hand, as many have pointed out, it is not clear that we should observe present bias in decisions about money—even if humans are present-biased. Utility is defined over consumption, so if subjects are able to borrow and save, intertemporal tradeoffs over dated money payments should depend on market interest rates, not individual preferences (Coller and Williams 1999). Experimental economists, in contrast, have long argued that experimental subjects "narrowly bracket" their decisions in the lab, viewing dated monetary payments as though they were a consumption plan, and numerous experimental studies have supported this view (Andersen et al. 2008; Rabin and Weizsacker 2009). However, recent evidence has called narrow bracketing assumption into question. For example, Augenblick et al. (2015) find evidence of present bias in effort tasks, but only limited evidence of present bias in decisions about money. Abdellaoui

[1] In other words, consumption that occurs t periods in the future is discounted by a factor δ^t, where $\delta \leq 1$ and does not vary over time. See Frederick et al. (2002) for an overview of the development of discounted utility model and its use in economics.

[2] In the discounted utility model, agents care more about immediate payments than about payments that occur k days in the future, but only as much as they care more about payments at time t than payments at time $t + k$.

[3] Within psychology, the most widely used model of present bias is the modified hyperbola (Kirby 1997). In that model, utility takes the form: $U(c_t) = \frac{1}{1+kt}u(c_t)$.

et al. (2018) find limited evidence of present bias for both consumption and money. Dean and Sautmann (2016) find that intertemporal tradeoffs in their lab-in-the-field experiment are associated with both expenditure shocks and savings, suggesting that narrow bracketing fails in their data. These results have sparked a lively debate, with some scholars arguing that choices in time preference experiments are driven primarily by liquidity constraints and interest rates outside the lab (Dean and Sautmann 2016; Epper 2015; Carvalho et al. 2016), while others maintain that there is little evidence that agents integrate moderately-sized monetary payments into their optimal lifetime consumption plan through smoothing and arbitrage (Halevy 2014, 2015).

Paralleling this rising chorus of theoretical objections, there is mounting concern that standard experimental designs used to measure time preferences may be confounded. For example, Frederick et al. (2002) point out that many experimental studies documenting present bias ask subjects to choose between smaller, immediate cash payments—which are typically given out at the end of the experimental session—and larger, delayed payments. If subjects are not sure that they will actually receive the later payment, or collecting delayed payments involves larger transaction costs, they may appear present-biased when in fact they are not (Halevy 2008; Andreoni and Sprenger 2012b; Gabaix and Laibson 2017). Another concern is that many experiments assume that utility is linear in money; such an assumption will lead to over-estimates of the degree of present bias if subjects are risk averse (Andersen et al. 2008).

In an attempt to address many of these methodological issues, Andreoni and Sprenger (2012a) introduced a novel experimental design—the convex time budget (CTB) experiment. In a CTB experiment, a subject divides an endowment between two time periods subject to a budget constraint and an interest rate that makes the delayed payment date relatively attractive. Because subjects are not restricted to the endpoints of the budget line, this method allows for separate estimation of the time preference parameters and the curvature of the utility function. Andreoni and Sprenger (2012a) conduct CTB experiments in a university lab setting that allows them to take a number of steps to equalize transaction costs and uncertainty across time periods. Importantly, they make same-day payments using the same technology as delayed payments (checks in campus mailboxes). After introducing such protocols, they find no evidence of present bias among university undergraduates, casting further doubt on the existence of present bias over money payments.[4]

One concern with several recent studies focused on equalizing transaction costs across immediate and delayed payments is that the steps taken to to do so also introduce a small front-end delay. For example, Andreoni and Sprenger (2012a) make "immediate" payments by placing a personal check in each subject's mailbox before the close of the business day.[5] Thus, "immediate" payments may not always be

[4] Harrison et al. (2013) have criticized the CTB task for relying on corner choices in identifying present bias. This criticism is orthogonal to our goal in the present paper, which is to compare the degree of present bias across immediate and end-of-day treatments. In addition, in contrast to previous findings, our subjects choose mostly interior allocations.

[5] Other studies in a similar vein involve even greater delays. For example, in Giné et al. (2017), the soonest payments occur 1 day after decisions are made. In Carvalho et al. (2016), checks are mailed on the day decisions are made, so they arrive at least 1 day later.

accessible immediately. If subjects do, in fact, have preferences consistent with the quasi-hyperbolic model, it is possible that they may view such almost-immediate payments as "later" rather than "now"—in which case, some of the recent failures to reject the discounted utility model may be attributable to the use of under-powered experimental tests.[6]

We test whether delaying payments until the end of the day attenuates present bias by conducting a series of convex time budget experiments at the Busara Center for Behavioral Economics in Nairobi, Kenya. We conducted two experimental treatments which differ in terms of payment timing. In our immediate payment treatment, all dated payments arrived at the time of day when experimental sessions concluded; hence, same-day payments arrived immediately after the experimental session. In our end-of-day payment treatment, all payments arrived near the close of the business day. In both treatments, all payments were made using Kenya's mobile money payment system, M-Pesa, which made it possible to send payments to participants in real-time. M-Pesa payments are widely accepted throughout Kenya, and could be converted to cash by walking into a shop across the street from the experimental lab. Same-day payments in the immediate payment treatment were delivered to participants through their phones as they left the experimental session—so, the earliest possible payments were truly immediate. Thus, we are able to equalize transaction costs and uncertainty by delivering both truly immediate and delayed payments through M-Pesa—while making "now" more immediately accessible than in many previous CTB experiments.

We implemented our CTB experiment using a user-friendly touchscreen computer interface that allowed us to collect a large data set of 48 CTB decisions from every subject—while working in a population that is substantially less affluent, educated, and elite than standard subject pools of students at top universities in the U.S. and Europe. This allows us to estimate preference parameters at the individual level and to explore the association between liquidity constraints and estimated preference parameters. The stakes in our experiment were large in terms of subjects' purchasing power: the median total payment was more than four times the median level of daily expenditure.[7] Thus, subjects had every incentive to think carefully about their decisions, and our design provides a powerful test of the extent of arbitrage between lab and non-lab savings vehicles.

We report three main findings. First, and most importantly, our results suggest a substantial degree of present bias over money in the immediate payment treatment, but little or no present bias when the earliest possible payments occur at the end of

[6] Augenblick et al. (2015) is an important exception: they deliver cash payments at the end of their experimental sessions, and report more limited evidence of present bias over money than over effort. However, their findings do suggest at least a modest degree of present bias over money. Specifically, their subjects allocate 38.1% (SE: 1.73) of the budget to the sooner payment date for monetary decisions not involving today, and 41% (SE: 1.34) for decisions involving today. The difference of 2.1 percentage points is marginally significant (p value 0.07) in their sample of 75 subjects. As discussed further below, the magnitude of the difference is quite similar to the difference of 2.8 percentage points observed in our study.

[7] Thaler (1981) first observed that subjects tend to appear more patient when making intertemporal tradeoffs involving larger stakes. More recently, Sun and Potters (2016) show that changes in stakes impact the estimated degree of impatience (i.e. the exponential discount factor) but not the degree of present bias. Daily expenditure is the best benchmark for stake size because income in our setting is lumpy.

the day. Our preferred empirical specifications indicate that delayed payments are discounted by between 7.6 and 9.8% relative to truly immediate payments (i.e. estimated β parameters for the immediate payment treatment range from 0.902 to 0.924), while immediate payments that arrive at the end of the day are discounted by less than 1% relative to future payments. We can reject the hypothesis that the degree of present bias is equal across treatments (p values 0.014 to 0.025 in our preferred specifications). Our central result is robust to a wide array of robustness checks, and we find a similar pattern when we estimate preference parameters at the individual level without making any assumptions about preference homogeneity or the distribution of preference parameters in our sample.

Our second main finding is that individual time preference parameters are not significantly related to measures of liquidity constraints, suggesting that such constraints are not a plausible alternative account of our findings. Moreover, subjects do not display any tendency to shift experimental payments toward days when they anticipate having limited cash-on-hand. Thus, it is highly unlikely that we are falsely ascribing to present bias patterns of behavior that are actually driven by liquidity constraints.

Third, we find that most subjects who are not liquidity-constrained do not engage in the sorts of arbitrage we would expect if they were integrating their experimental payments into an optimal forward-looking consumption and savings plan. The overwhelming majority of subjects who hold substantial liquid savings sometimes choose interior allocations in CTB decision problems which offer gross interest rates over 100% (over a 4 week time horizon)—well above those available through the credit market.[8] Thus, they do not fully exploit the investment opportunities offered by the experiment, even though doing so would increase the net present value of their income (and therefore consumption) stream.

Taken together, our results demonstrate that present bias over money is not simply an artifact of experimental design flaws in previous studies; we find strong evidence of present bias in an experiment that controls for risk aversion, using protocols that equalize transaction costs and payment modalities across all possible payment dates. However, present bias does appear to depend on immediacy: it is nearly eliminated when the earliest possible payments do not arrive until the end of the day. Our estimates of present bias at the end of the day are similar to those found

[8] In our sample, 59.7% of chosen allocations are interior, and only 11.9% of subjects always choose corner solutions (which would be consistent with either risk neutrality or arbitrage). The pattern of behavior among our adult subjects stands in marked contrast to the patterns observed in several recent studies of university students in the United States and Europe. For example, Augenblick et al. (2015) report that only 14% of CTB decisions over money payouts are interior and 61% of their student subjects (in the U.S.) always choose corner solutions, while Sun and Potters (2016) report that 30% of chosen monetary allocations are interior and 37% of student subjects (in the Netherlands) always choose corner solutions. Interestingly, our results line up with those of Giné et al. (2017), who report that only 16.5% of CTB decisions by Malawian farmers are at corners (they do not report the proportion of subjects who never choose an interior allocation). Though comparisons across studies are inherently speculative, the pattern of evidence appears to suggest that adult subjects in low-income countries (Kenya and Malawi) are less likely to behave in a manner consistent with arbitrage than student subjects in wealthy countries (the U.S. and the Netherlands).

in Western settings, suggesting that cross-country differences in preferences cannot account for our findings (Falk et al. 2018). Our study also provides clear evidence that subjects—specifically, a diverse sample of adults in a lower middle income country—are not arbitraging between lab and outside savings vehicles; we also find no evidence that choices in our experiment are driven by liquidity constraints. Thus, our the results support the view that individual choices in time preference experiments are driven by time preferences, not market interest rates.

Our findings help to reconcile several disparate results in the literature. As discussed above, our results highlight the importance of immediacy. Of eight existing papers that test for present bias over money using convex time budget experiments, five introduce a small front-end delay (or, at least, the possibility of a small front end delay) before the earliest payments (Andreoni and Sprenger 2012a; Clot et al. 2015; Carvalho et al. 2016; Giné et al. 2017; Sun and Potters 2016). Only Carvalho et al. (2016) report finding present bias, and they only detect present bias when decisions are made prior to a subject's payday. In contrast, two of the three CTB experiments that offered immediate payments report substantial present bias over money (Luhrmann et al. 2013 and our study). Luhrmann et al. (2013) suggest that their results differ from those of Andreoni and Sprenger (2012a) because they focus on teenagers rather than university students. However, when viewed together with our results and juxtaposed against the CTB papers involving a small front-end delay, it seems reasonable to conclude that present bias may only be relevant in decisions involving immediate payments. Indeed, though Augenblick et al. (2015) find substantially less present bias over money than over effort, their results are consistent with a modest amount of present bias over money. Thus, the aggregate evidence suggests that present bias is quite sensitive to payment timing: even minor delays may mute the extent to which tradeoffs are perceived as "now" versus "later." This gives firms and policymakers considerable power over consumers, since minor modifications to contract structure (that have almost no consequences for firms) can lead to substantial changes in revealed preferences.

The remainder of the paper is structured as follows. Section 2 describes the design and implementation of the study. Section 3 presents our theoretical framework and derives testable predictions. Section 4 presents our exp rimental main results. Section 5 discusses the relationship between our work and other recent time preference experiments, and Sect. 6 concludes.

2 Experimental Design and Procedures

2.1 Experimental Design

We employ the convex time budget (CTB) design first used by Andreoni and Sprenger (2012a). In a CTB experiment, each subject divides a budget m between two payment dates subject to the early-valued budget constraint:

$$c_t + \frac{c_{t+k}}{(1+r)} = m. \tag{1}$$

In this framework, r is the interest rate; t denotes the front-end delay, the number of days between the experiment and the earlier payment date; and k denotes the delay between the earlier and later payment dates. Subjects in our experiment faced a total of 48 CTB decision problems, each of which was presented using a user-friendly touchscreen computer interface. This design generates an extremely rich data set and allows us to estimate preference parameters at the individual level. The CTB decisions included in our experiment were organized into eight sets of six decision problems. The earlier and later payment dates were fixed within each set. Within each decision set, the maximum earlier payment was fixed at either 400 or 600 Kenyan shillings.[9] The eight decision sets were presented in a random order.

Within each decision, the maximum later payment depended on the gross interest rate, $1 + r$: reducing the earlier payment (c_t) by 1 shilling meant increasing the later payment (c_{t+k}) by $1 + r$ shillings. Within each set of decisions, subjects faced six gross interest rates: 1.1, 1.25, 1.75, 2, 3, and 4. Gross interest rates always appeared in increasing order within a decision set to minimize the potential for confusion. The Online Appendix lists the front-end delay, delay between payments, budget size, and gross interest rate for each of the 48 CTB decisions included in our experiment.[10]

At the end of the experiment, one decision problem was randomly chosen to determine final payments. This randomization was done separately for each subject, guaranteeing that all information on the timing and size of experimental payments remained private. Randomization thus occurred within session, not across sessions. In addition to their payments from the experiment, subjects received a fixed show-up fee which was evenly divided between the earlier and later payment dates—so every subject, including those who chose corner solutions, received two dated payments.[11] We describe the procedures used to deliver payments to subjects in detail below.

[9] These budgets are equivalent to approximately 4.08 and 6.12 USD, respectively. These endowments are large in purchasing power terms: the median level of daily expenditures in our sample is 146 Kenyan shillings (1.49 USD)

[10] At the end of the CTB portion of the experiment, subjects completed a standard Multiple Price List (MPL) task that included 24 decision problems. One of the 72 decision problems was randomly selected to determine experimental payouts. See Charness et al. (2016) for discussion of the consequences of paying for a single randomly-selected decision problem.

[11] As Andreoni and Sprenger (2012a) discuss, splitting the show-up fee across the two payment dates is critical to the research design. Haushofer (2014) presents a theoretical model suggesting that a mental cost of keeping track of time-dated payments may act as an additional (cognitive) transaction cost, pushing subjects toward corner solutions and immediate payments when the show-up fee is paid (in its entirety) on the day of the experiment. However, this cost would apply to both payment dates in our setting because half of the show-up fee is paid on each payment date. Thus, if any transaction cost enters as an additively separable (from money/consumption utility) term in the utility function, it should not impact allocation decisions at all (because choosing a corner solution would not reduce the amount received on either date to 0, so subjects would still incur the transaction costs associated with each of the two dated payments). Alternatively, if the transaction cost enters as a reduction in money utility that is larger for delayed payments, allocations to the earlier payment date should be lower when the earliest payments are immediate, since utility is (weakly) concave. Taken together with stated beliefs about the likelihood that payments will arrive on time (and subjects' experience with the Busara lab's reliability), it is quite unlikely that differential transaction could explain behavior in our experiment.

2.2 Experimental Procedures

The experiment was conducted at the Busara Center for Behavioral Economics in Nairobi, Kenya. Subjects were drawn from two of Nairobi's informal settlements, Kawangware and Kibera. Our sample includes data from 494 adult subjects. Summary statistics on the subjects in our sample are reported in Table A1 of the Online Appendix.

Experimental sessions were conducted in a dedicated computer lab at the Busara Center.[12] Instructions were presented orally in Swahili, one of Kenya's official languages and a local *lingua franca*. Our user-friendly touchscreen interface was programmed using z-tree (Fischbacher 2007), and was intended to be easily comprehensible by subjects with limited formal education. Full experimental instructions and screenshots of the computer interface are included in the Online Appendix.

At the end of the experiment, the computer randomly selected one decision for payment. Payments from the chosen decision were added to a show-up fee, which was divided evenly between the earlier and later payment dates (from the decision that was chosen to determine the final payment). Thus, all subjects received two dated payments. Online Appendix Figure A1 summarizes the structure of each experimental session.

All payments in our experiment—including payments made on the day of the experiment—were made using the M-Pesa mobile money technology. M-Pesa is a money transfer service operated by Kenya's largest mobile phone company, Safaricom. Users can send and receive transfers and make direct payments to firms using their phones, and they can also withdraw cash from their M-Pesa accounts at over 80,000 M-Pesa agents throughout the country. All subjects in our experiment had active M-Pesa accounts, and all had received transfers from the Busara Center via M-Pesa prior to the experiment. So, transaction costs for immediate and delayed payments were equalized—all payments were sent from a trusted source (the Busara Center) via the familiar M-Pesa technology.

2.3 Experimental Treatments

To test whether small front-end delays of less than 1 day reduce present bias, we conducted two experimental treatments which differ in terms of payment timing. In the immediate payment treatment, same-day payments were guaranteed to arrive no more than two hours after the start of the experimental session. Subjects receiving payments on the day of the experiment would typically receive them before they departed from the Busara Center, while they were completing the post-experiment survey. Thus, in the immediate payment treatments, same day payments were truly

[12] No experimental sessions were held on Fridays or weekends to avoid any potential end-of-week effects. When considering a potential date for a session, we verified that no payment dates associated with that (potential) session fell on holidays or other days likely to lead to foreseeable changes in the desire for cash on hand (for example, the day when school fees are due). We test whether later payment dates are associated with individual needs for ready cash in Sect. 5.2.

immediate in the sense that they could be spent as soon as the experiment was over.[13,14]

In the end-of-day payment treatment, payments were guaranteed to arrive by 6 o'clock in the evening. Thus, same day payments arrived on the day of the experiment, but were not immediately accessible. In both treatments, all sessions were held in the mornings, so subjects assigned to the end-of-day treatments had to wait at least four hours before they could access their experimental payments. Within each session, payments arrived at the same time on every possible payment date, so the only difference between the immediate payment treatment and the end-of-day payment treatment was that payments arrived late in the afternoon, several hours after the time of the experiment, in the end-of-day payment condition.[15,16]

3 Theoretical Framework

Each subject in our experiment divides a budget of $m > 0$ between two accounts: one associated with an earlier payment date ($t \geq 0$ days in the future) and one associated with a later payment date ($t + k > 0$ days in the future). Following Laibson (1997) and O'Donoghue and Rabin (1999), we consider a subject i who maximizes her (additively separable) utility which can be represented by

$$u(c_t, c_{t+k}) = u(c_t + \omega) + \beta\delta^k u\left(c_{t+k} + \omega\right) \text{if } t = 0 \tag{2}$$

and

$$u(c_t, c_{t+k}) = u(c_t + \omega) + \delta^k u\left(c_{t+k} + \omega\right) \text{if } t \neq 0 \tag{3}$$

subject to the budget constraint

$$c_t + \frac{c_{t+k}}{(1 + r)} = m. \tag{4}$$

[13] M-Pesa withdrawals can be made at any one of many M-Pesa agents in Nairobi, typically located in shops or kiosks. There are numerous M-Pesa agents in the immediate vicinity of the Busara Center, and also in the informal settlements where participants live. For example, as of 2011, Safaricom reported that there were 150 active M-Pesa agents in the single square-mile Kibera slum neighborhood, and there were at least 100 in the smaller Kawangware neighborhood. Many of these agents are open late into the night. Our payments were made before 6:00 PM, ensuring that participants were able to withdraw the money on the day of the study if they wished.

[14] Because subjects had experience receiving mobile money payments from the Busara Center via M-Pesa, there is little reason to be concerned that they doubted that their payments would arrive on time. When asked (at the end of the experiment) whether they thought that both of their experimental payments would arrive on time, 98% of subjects answered in the affirmative.

[15] In Table A2 of the Online Appendix, we show that observable characteristics are comparable across experimental treatments. The notable exception is that subjects assigned to the immediate payment treatment appear less likely to be liquidity-constrained than those assigned to the end-of-day payment treatment. Of course, if behaviors that appear present-biased were actually driven by liquidity constraints, this imbalance would predict greater present bias in the end-of-day payment treatment than in the immediate payment treatment.

[16] Sessions were conducted over two years: 2015 and 2016. Immediate payment sessions were conducted in both 2015 and 2016, while all end-of-day payment sessions were held in 2016.

In this framework, δ is the per-period exponential discount factor. β indicates the degree of present bias—i.e. the extent to which all future payoffs are discounted relative to immediate payoffs. ω denotes background consumption (e.g. the show-up fee or income from outside the experiment); this will be equal to 0 if subjects narrowly bracket their decisions in the experiment (Rabin and Weizsacker 2009). When individual preferences are dynamically consistent (i.e. when $\beta = 1$), the optimal c_t^* depends on the size of the budget (m), the gross interest rate (r), and the delay between payments (k)—but not on the front end delay (t). However, the exponential discounting model is the only model that generates dynamically consistent choices. In the equation above, when $\beta < 1$, the optimal c_t^* depends on t: the optimal allocation to the earlier period is higher when $t = 0$ than for all $t > 0$. Such changes in the optimal c_t^* as t changes are termed static preference reversals.

If we assume that utility takes the constant relative risk aversion (CRRA) form such that

$$u(c_t) = \frac{c_t^{1-\rho}}{1-\rho}, \tag{5}$$

then we can solve for the demand function for c_t, and the parameters β, δ, and ρ can be estimated by non-linear least squares or maximum likelihood—so we can test the hypothesis that $\beta < 1$ directly in a structural framework. As we discuss further below, we also estimate the effect of our end-of-day payment treatment on the estimated β parameter; this allows us to assess the extent to which present bias is attenuated when payments are not immediate.

4 Analysis

4.1 Comprehension and Consistency

An important question in all preference elicitation experiments is whether subjects make coherent choices that are consistent with utility maximization. There is ample evidence that even university student subjects implement their choices with error (Hey and Orme 1994; Andreoni and Miller 2002; Choi et al. 2007a; Fisman et al. 2007). However, though individual choices are often noisy, studies of both university students and general population samples in wealthy countries show that most experimental subjects make decisions that are far more consistent that we would expect to occur by chance (Choi et al. 2014; Fisman et al. 2017). Nevertheless, concerns about comprehension and consistency are particularly salient in our context because our subjects have less education and experience with computers than typical experimental subject pools composed of university students.

We take two approaches to assessing the consistency of subjects' choices. First, we test choices for consistency with the Generalized Axiom of Revealed Preference (GARP). GARP provides a direct test of whether individual choices can be rationalized by a utility function that is continuous, increasing, concave, and piecewise linear (Afriat 1967; Varian 1982, 1983). As we discuss in detail below, the majority of subjects in our experiment do not violate GARP. However, because of the small number of intersecting budget lines in our experiment, the power of our GARP test is somewhat limited. We therefore adopt a second approach: testing the extent of adherence to the law of demand.[17] Taken together, both approaches suggest that subjects understood the experiment and made consistent decisions that can be viewed through the lens of utility maximization.

4.1.1 Rationality

One of the most important questions one can ask about individual decision data is whether choices are consistent with utility maximization. When budgets are linear, revealed preference theory offers a direct test of rationality: choices can be rationalized by a utility function that is well-behaved (in the sense of being continuous, increasing, concave, and piecewise linear) if and only if they satisfy GARP (Afriat 1967). By varying both the (early-valued) budget size and the gross interest rate for fixed pairs of earlier and later payment dates (i.e. fixed values of t and $t+k$), we confront subjects with sets of intersecting budget lines that create the possibility of violating GARP. Specifically, our experiment includes two sets of 12 intersecting budget lines in which the earlier and later payment dates are fixed but the budget sizes and interest rates vary across decision problems. Our tests of consistency with GARP are based on choices in these 24 decision problems.

To assess the power of our GARP test, we follow the standard approach, which builds on Becker (1962) and Bronars (1987), generating a population of 1000 simulated subjects who choose points at random from each budget line (according to a uniform distribution).[18] 86.9% of these simulated subjects violate GARP at least once, suggesting that our revealed preference test does, in fact, have a reasonable level of power. The median number of violations in the sample of simulated subjects is 8. In contrast, 61.3% of subjects in our experiment never violate GARP, and only 19.2% have 8 or more violations. Figure 1 presents histograms of the distributions of GARP violations in our actual and simulated samples. Though the power of our GARP test is lower than in some recent experiments (cf. Fisman et al. 2015), the evidence suggests that our subjects are substantially closer to consistency with utility maximization than could occur at random.

4.1.2 Adherence to the Law of Demand

To further gauge the extent to which subjects in our experiment made meaningful and consistent choices, we follow Giné et al. (2017) in examining "basic

[17] In CTB experiments where the budget size is denominated in terms of the later payment date (so that the maximum later payment is fixed as the interest rate changes), this is often referred to as "monotonicity" (Andreoni and Sprenger 2012a) or "demand monotonicity" (Chakraborty et al. 2017). We adopt the terminology used by Giné et al. (2017) since, like them, we use an early-valued convex time budget.

[18] See Choi et al. (2007b) for discussion.

consistency"—a measure of adherence to the law of demand. The idea underlying basic consistency is that, for a fixed t and k (i.e. fixed earlier and later payment dates), an increase in the gross interest rate is equivalent to a decrease in the price of consumption in the later period. So, if we consider two interest rates, r' and r'' such that $r'' > r'$, the amount allocated to the later period should be at least as large under r'' as under r'.

Subjects in our experiment made 8 sets of 6 CTB decisions. Within each set, the budget size (i.e. the maximum earlier payment), t, and k were fixed. Each set of decisions included 6 gross interest rates: 1.1, 1.25, 1.75, 2, 3, and 4. There are therefore 15 possible pairs of interest rates in each set of decisions. For each pair, we generate an indicator for a basic consistency violation that is equal to 1 if the allocation to the later account is higher under the lower of the two interest rates. We then calculate individual-level frequency of such violations. One minus the frequency of basic consistency violations provides an index of the level of adherence to the law of demand. The median rate of basic consistency is 0.93, suggesting that subjects understood the experiment and were able to implement purposeful choices using the computer interface.

For comparison, we again follow the approach suggested by Bronars (1987), generating a population of simulated subjects who choose points from each budget line randomly (according to a uniform distribution). The median rate of basic consistency among simulated subjects is only 0.72, and only 6.2% have basic consistency indices of at least 0.8 (versus 76.5% of actual subjects). Figure 2 compares the distribution of the basic consistency index in our (actual) sample to the simulated distribution. It is clear that a large majority of subjects make choices that are much more consistent than what would occur by chance.

4.2 Summarizing Individual Choices in the Experiment

With consistency established, we now examine the intertemporal tradeoffs made by subjects in our experiment. Figure 3 plots the fraction of the (early-valued) budget allocated to the earlier payment date as a function of front-end delay and the gross interest rate.[19] Panel A summarizes subjects' choices in the immediate payment treatment (where the earliest possible payment occurred immediately after the experiment). The figure suggests some degree of present bias: when the front-end delay is 0, subjects allocate more to the early payment date. The effect is relatively modest, however. Subjects in the immediate payment treatment allocate an average of 47.7% of their early-valued budgets to the earlier payment date when there is no front-end delay, versus 44.9% when the front-end delay is either two or four weeks. Although the effect is fairly small, it is consistent: subjects allocate more to the earlier payment date when early payments are immediate across the full range of interest rates in the experiment.

[19] CDFs of the budget fraction allocated to the earlier payment date are presented in Online Appendix Figure A2.

Panel B of Fig. 3 presents results from the end-of-day payment treatment, when the earliest possible payments arrived late in the afternoon on the day of the experimental session. Here, we observe little if any evidence of present bias: the budget fraction allocated to the earlier payment when the front-end delay is 0 days (but several hours, since all payments in these treatments occurred at the end of the day) is virtually identical to the fraction allocated to the earlier payment date when the front-end delay is longer. Subjects allocate an average of 43.6% of their early-valued budgets to the earlier payment date when the front-end delay is 0, versus 42.9% when the front-end delay is more than 1 day. Thus, the aggregate pattern suggests that present bias is almost entirely eliminated when immediate payments are delayed until several hours after the experimental session.

Next, we explore these patterns in a regression framework. We regress the budget fraction allocated to the earlier payment date on an indicator for decisions where the front-end delay is 0, an indicator for the end-of-day payment treatment, and an interaction between the two. In all specifications, we also include controls for the budget size, the interest rate, the delay between payments, and the time of day when the experimental session was held. We report OLS specifications as well as Tobit specifications that adjust for censoring of the dependent variable at 0 and 1. Standard errors are clustered by session.

Results are reported in Table 1. In all specifications, subjects allocate significantly more money to the earlier payment date when the front-end delay is 0—suggesting some degree of present bias in the immediate payment treatment. Point estimates indicate that subjects allocated 2.9 to 4.3 percentage points more of their early-valued budgets to the earlier payment date when early payments occurred immediately after the experiment (i.e. when the front-end delay was 0 in the immediate payment treatment). Though the effect is relatively modest, it is significant at the 99% confidence level in both the OLS and the Tobit specifications (p values < 0.001).[20]

In contrast, we can never reject the hypothesis that allocation decisions do not depend on the degree of front-end delay in the end-of-day payment treatments. The interaction between the indicators for the end-of-day payment treatment and for decisions involving no front-end delay is consistently negative and significant; this indicates that the treatment effect of the front-end delay of 0 is smaller in the end-of-day payment treatments than in the immediate payment treatments. Tests of the overall impact of same-day payments on the allocation to the earlier payment date consistently fail to reject the null for the end-of-day payment treatment (p values 0.462 and 0.693). Thus, our reduced form results suggest that delaying payments by a few hours all but eliminates present bias.

4.3 Estimating the Degree of Present Bias

Next, we test for the presence of present bias by estimating β directly in a structural framework. We begin by estimating the β, δ, and ρ parameters via non-linear least

[20] As discussed above, the magnitude of this reduced form effect is comparable to that observed by Augenblick et al. (2015).

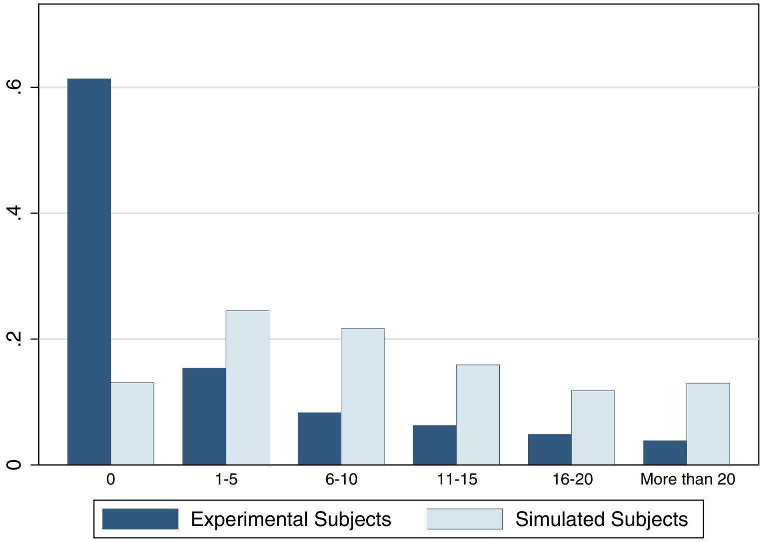

Fig. 1 Frequency of GARP violations for actual versus simulated subjects. The figure reports the total number of violations of the Generalize Axiom of Revealed Preference for experimental subjects and for 1000 thousand simulated subjects who randomize uniformly on each budget line. Violations are calculated for the two pairs of (earlier and later) dates that were presented at both the 400 shilling (early-valued) budget size and the 600 shilling (early-valued) budget size (i.e. decision sets 1, 3, 4, and 7 from the CTB experiment), creating intersecting budget lines that allow for revealed preference tests

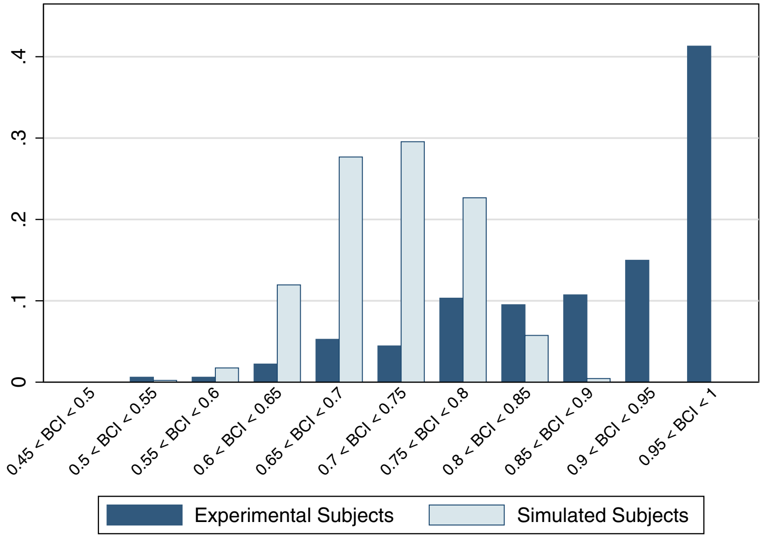

Fig. 2 The Basic Consistency Index for actual versus simulated subjects. To calculate the Basic Consistency Index (BCI), we consider all pairs of interest rates, r' and r'', such that $r'' > r'$; a subject's decisions satisfy basic consistency if the amount allocated to the later period is at least as large under r'' as under r'. The BCI is fraction of all possible pairs of choices that satisfy this property

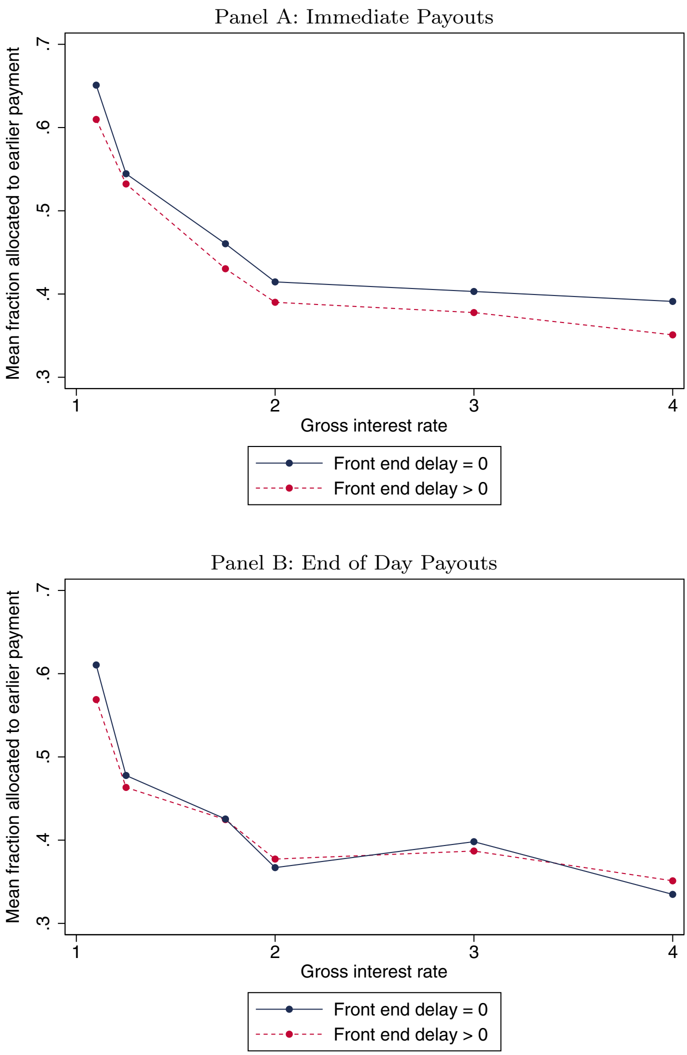

Fig. 3 Fraction of budget allocated to earlier account, by interest rate

squares (NLS), allowing all three parameters to differ across treatments.[21] We report three specifications which take different approaches to the background consumption parameter, ω. First, we impose the assumption that subjects narrowly bracket their decisions in the experiment by setting ω equal to 0. We then allow ω to vary across subjects by proxying for background consumption with self-reported average daily

[21] To facilitate comparisons of parameter magnitudes across treatment (without necessitating an unduly large number of digits), we report weekly discount factors throughout the analysis.

Table 1 The impact of front-end delays on allocation decisions

Specification	OLS (1)	OLS (2)
Front-end delay = 0 days	0.029***	0.043***
	(0.006)	(0.01)
End-of-day payment treatment	– 0.014	– 0.027
	(0.021)	(0.033)
Front-end delay = 0 days × end-of-day treatment	– 0.022**	– 0.037**
	(0.011)	(0.018)
H_0: no impact of front-end-delay = 0 days in end-of-day treatment	0.462	0.692
Observations	23,712	23,712
Subjects	494	494

Robust standard errors clustered at the session level. ***, **, and * indicate significance at the 99, 95, and 90 percent levels, respectively. The dependent variable in all specifications is the fraction of the early-valued budget allocated to the earlier payment date. The Tobit regression (in Column 2) adjusts for censoring of the dependent variable at 0 and 1. All regressions include controls for the size of the early-valued budget, the interest rate, the delay between payment dates, and the experimental session time

expenditure, ω_i. Finally, we estimate ω as one of the model parameters. Results are reported in Table 2.

Though parameter estimates differ slightly across specifications, they paint a consistent picture. The estimated β parameters are significantly less than 1 in the immediate payment treatment (all p values < 0.001), indicating that subjects' choices are present-biased. The estimates of β range from 0.902 to 0.924. In the end-of-day payment treatment, we do not observe a statistically significant degree of present bias. The estimates of β are higher, ranging from 0.982 to 0.992—suggesting at most an extremely modest amount of present bias when payments are made several hours after the experimental session. In all specifications, we can reject the hypothesis that the degree of present bias is equal in the immediate and end-of-day payment treatments (p values range from 0.014 to 0.025 across specifications).

Turning to the estimated δ parameters, we find that subjects in both the immediate and the end-of-day payment treatments are extremely impatient. Estimated weekly discount factors range from 0.942 to 0.950 in the immediate payment treatment, versus 0.961 to 0.972 in the end-of-day payment treatment. All are significantly different from 1 at at least the 95% confidence level. These estimates suggest that payments 1 year in the future are discounted by between 77 and 96%.[22]

The estimates of ρ and ω are broadly comparable in the immediate and end-of-day payment sessions. The estimates of ρ are (unsurprisingly) higher in the specifications

[22] These levels of impatience are higher than those observed in studies in Western populations; e.g., Andreoni and Sprenger (2012a) report yearly discounting between 30 and 38%. This difference is in line with recent evidence showing higher levels of impatience in Sub-Saharan Africa compared to North America and Europe (Falk et al. 2018).

Table 2 NLS estimates of model parameters in immediate versus end-of-day payment treatments

Specification	NLS (1)	NLS (2)	NLS (3)
$\beta_{immediate}$	0.902***	0.920***	0.924***
	(0.014)	(0.012)	(0.012)
β_{eod}	0.982***	0.990***	0.992***
	(0.029)	(0.024)	(0.026)
$\delta_{immediate}$	0.950***	0.942***	0.942***
	(0.008)	(0.007)	(0.006)
δ_{eod}	0.972***	0.965***	0.961***
	(0.012)	(0.009)	(0.010)
$\rho_{immediate}$	0.581***	0.887***	0.945***
	(0.021)	(0.031)	(0.055)
ρ_{eod}	0.619***	0.920***	0.904***
	(0.026)	(0.042)	(0.037)
$\omega_{immediate}$	0	ω_i	248.477***
	–	–	(38.541)
ω_{eod}	0	ω_i	187.559***
	–	–	(29.205)
H_0: $\beta_{immediate} = 1$	0.000	0.000	0.000
H_0: $\beta_{eod} = 1$	0.532	0.666	0.749
H_0: $\beta_{immediate} = \beta_{eod}$	0.017	0.014	0.025
H_0: $\delta_{immediate} = 1$	0.000	0.000	0.000
H_0: $\delta_{eod} = 1$	0.032	0.001	0.000
H_0: $\delta_{immediate} = \delta_{eod}$	0.132	0.050	0.106
H_0: $\rho_{immediate} = \rho_{eod}$	0.253	0.523	0.544
H_0: $\omega_{immediate} = \omega_{eod}$			0.215
Observations	23,712	23,712	23,712
Subjects	494	494	494

Robust standard errors clustered at the session level. ***, **, and * indicate significance at the 99, 95, and 90 percent levels, respectively. ω_i indicates self-reported average daily expenditure, which varies across subjects

that allow for positive background consumption (Columns 2 and 3 of Table 2), but do not differ substantially across experimental treatments. Estimates suggest moderate risk aversion, with estimated values of ρ between 0.5 and 1 in all specifications.[23] Estimating the background consumption parameter, ω, structurally and allowing it to vary across experimental treatments (in Column 3 of Table 3) generates results that are quite similar to those obtained by using self-reported daily expenditures as a

[23] Though we take an entirely different approach to preference elicitation, our results are similar to estimates of risk aversion in broadly comparable field populations (cf. Harrison et al. 2010; Jakiela and Ozier 2016).

Table 3 OLS Regressions of Individual-Level Parameter Estimates

Estimated parameter	β_i	δ_i	ρ_i	β_i	δ_i	ρ_i
Specification	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
End of day payout treatment	0.061**	− 0.009	0.22	0.074***	− 0.011	0.288
	(0.029)	(0.02)	(0.333)	(0.027)	(0.023)	(0.339)
Female	–	–	–	0.022	0.026	0.671*
				(0.03)	(0.024)	(0.342)
Age	–	–	–	0.003	0.002*	0.011
				(0.002)	(0.001)	(0.018)
Completed primary school	–	–	–	0.074	0.133	− 1.153
				(0.12)	(0.085)	(0.819)
Completed secondary school	–	–	–	− 0.073**	− 0.024	0.111
				(0.036)	(0.025)	(0.307)
Married or cohabitating	–	–	–	0.06*	0.013	0.175
				(0.032)	(0.025)	(0.255)
Has a bank account	–	–	–	− 0.029	0.017	− 0.321
				(0.033)	(0.028)	(0.362)
Has 1000 KSH in a bank account	–	–	–	− 0.012	− 0.034	0.33
				(0.042)	(0.031)	(0.438)
Average daily expenditure (in shillings)	–	–	–	− 0.0002	− 0.00002	0.004**
				(0.0001)	(0.0001)	(0.001)
Describes self as very patient	–	–	–	0.037	0.005	0.049
				(0.039)	(0.024)	(0.299)
Observations	477	477	477	446	446	446
R^2	0.007	0.0003	0.001	0.051	0.027	0.046

Robust standard errors clustered at the session level. ***, **, and * indicate significance at the 99, 95, and 90 percent levels, respectively. The top 5% of parameter estimates are winsorized so that results are not driven by outliers

measure of background consumption (in Column 2 of Table 3). We do not find evidence that ω differs significantly across treatments.

4.3.1 Robustness Checks

We report a range of robustness checks in the Online Appendix. First, we estimate β, δ, and ρ parameters for each experimental treatment via maximum likelihood, assuming an additively separable, normally distributed error term (Online Appendix Table A3). We also estimate model parameters via two-limit Tobit estimation based on the tangency condition characterizing the optimal interior allocation (Online Appendix Table A4). Next, we replicate our NLS estimation in sub-samples of subjects who showed high levels of comprehension and consistency: those subjects who never violated GARP and those with basic consistency indices above 0.85 (Online Appendix Tables A5 and A6). As a final robustness check, we also report

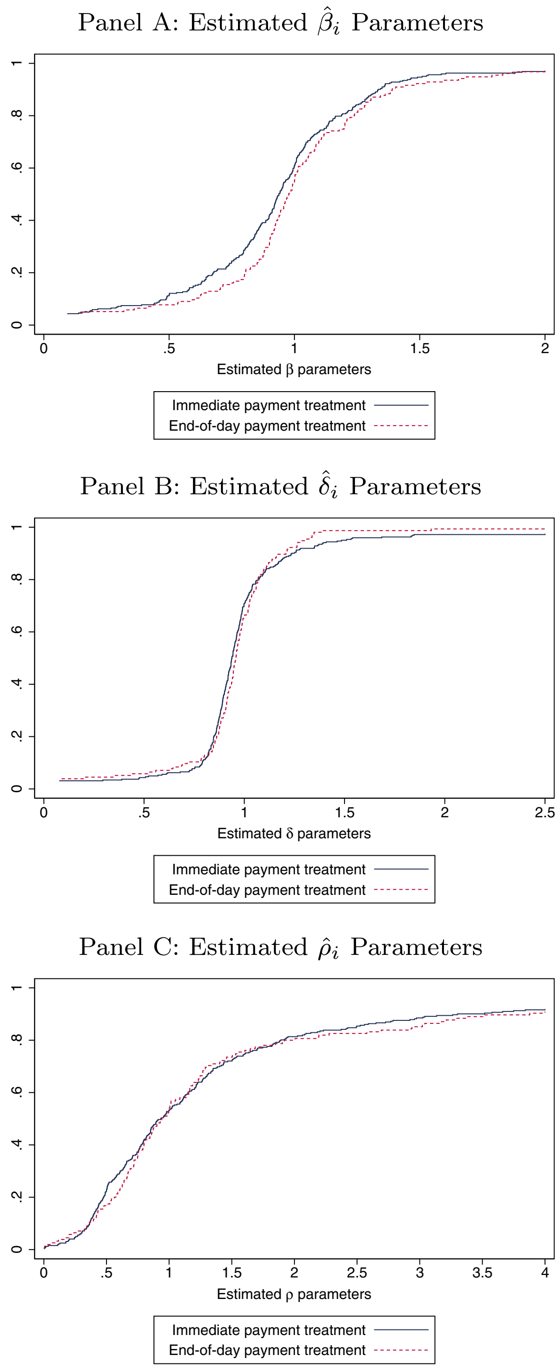

Fig. 4 Individual-level parameter estimates, by treatment

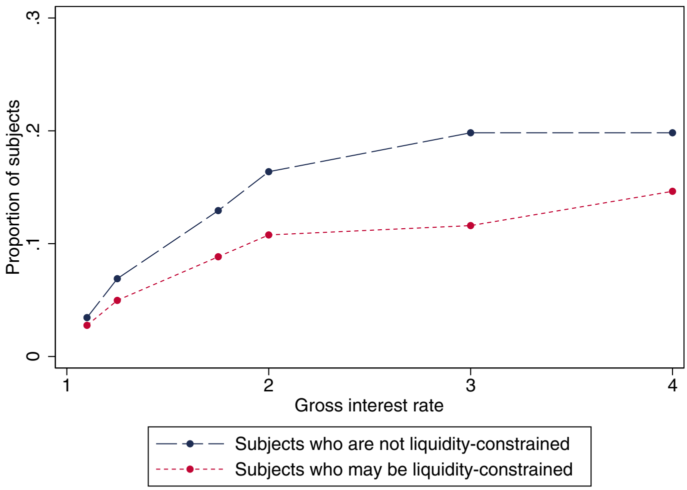

Fig. 5 The frequency of behavior consistent with arbitrage. The figure plots the proportion of subjects who always allocate their entire endowment to the later payment date at a given gross interest rate. Potentially liquidity-constrained subjects are those who do not have at least 1000 Kenyan shillings (USD 10.20) in a bank savings account

NLS estimates of β and δ derived under the assumption that utility takes the constant absolute risk aversion (CARA) form (Online Appendix Table A7). Our results hold across this range of specifications and functional form assumptions: choices in our immediate payment treatment provide strong evidence of present bias over money, but present bias appears to be attenuated substantially when payments occur at the end of the day.

4.4 Individual-Level Analysis

We next examine decisions in our experiment at the individual level. We estimate subject-level $\hat{\beta}_i$, $\hat{\delta}_i$, and $\hat{\rho}_i$ parameters via non-linear least squares while controlling for self-reported background consumption (average daily expenditure).[24] CDFs of the estimated individual-level $\hat{\beta}_i$, $\hat{\delta}_i$, and $\hat{\rho}_i$ parameters are presented in Fig. 4.[25] The distribution of individual-level $\hat{\beta}_i$ parameters in the immediate payment treatment is consistently to the left of the distribution the end-of-day payment treatment, suggesting greater present bias when payments are truly immediate. Indeed, the median individual-level $\hat{\beta}_i$ in the immediate payment treatment is 0.938, versus 0.978 in the

[24] We are unable to estimate individual parameters for 17 of our 494 subjects. 6 subjects always allocated their entire endowment to the earlier payment date, and 9 always allocated their entire endowment to the later payment date. Estimation does not converge for 2 of the remaining subjects.

[25] Like other recent studies of individual preferences (cf. Choi et al. 2007b; Fisman et al. 2007, 2017; Andersen et al. 2008), we observe tremendous individual heterogeneity in preferences, much of which is not explained by demographic and socioeconomic characteristics. The 5*th* percentile of $\hat{\beta}_i$ is 0.164, and the 95*th* percentile is 1.598. The 5*th* percentile of $\hat{\delta}_i$ is 0.473, and the 95*th* percentile is 1.388. The 5*th* percentile of $\hat{\rho}_i$ is 0.243, and the 95*th* percentile is 17.878.

end-of-day payment treatment, and we can reject the hypothesis that the medians are equal in the two treatments (p value 0.059). 61% of subjects in the immediate payment treatment have estimated $\hat{\beta}_i$ parameters below 1, versus 55% of those in the end-of-day payment treatment.[26] Thus, subjects who show some degree of present bias outnumber those who tend toward future bias by a wide margin, though present bias is far from universal.

Next, we examine our estimated $\hat{\delta}_i$ parameters. The median $\hat{\delta}_i$ in the immediate payment treatment is 0.938, versus 0.958 in the end-of-day payment treatment. These weekly discount factors suggest that subjects in the immediate payment treatment discount payments 1 year in the future by 96%, versus 89% for the end-of-day payment treatment. Our results also suggest that present bias and impatience are positively correlated (Spearman's $\rho = 0.176$).

To further test whether delaying experimental payments attenuates present bias, we report OLS regressions of the estimated individual-level $\hat{\beta}_i$, $\hat{\delta}_i$, and $\hat{\rho}_i$ parameters on an indicator for the end-of-day payment treatment (Table 3). When no controls are included, coefficient estimates suggest that $\hat{\beta}_i$ is significantly higher in the immediate payment treatment than in the end-of-day payment treatment (p value 0.044).[27] In contrast, assignment to the end-of-day payment treatment does not impact patience ($\hat{\delta}_i$) or risk aversion ($\hat{\rho}_i$). Results are similar when we include controls for individual characteristics such as age, gender, and education level (in Columns 4 through 6 of Table 3). Thus, though we observe substantial heterogeneity in individual preference parameters, our individual-level analysis confirms the main conclusion of our aggregate analysis: present bias is reduced when payments are not made immediately after the experimental session.

5 Discussion

Over the last few years, several theoretically sophisticated, technically rigorous experiments have sparked a lively debate about the use of lab experimental methods to measure intertemporal tradeoffs.[28] This body of work raises two critical questions. First, do lab experiments with money payments measure time preferences? Utility is defined over consumption, not money, and subjects have access to a range of credit products. So, sophisticated subjects may treat dated experimental payments as a(nother) form of credit, integrating their (present-discounted) experimental income into an optimal forward-looking consumption plan—in which case, choices in the lab will reflect market interest rates and (perhaps) individual liquidity constraints,

[26] We cannot reject the hypothesis that the proportion of subjects with estimated $\hat{\beta}_i$ parameters below 1 is equal in the immediate and end-of-day payment treatments (p value 0.211).

[27] We can reject the hypothesis that the average individual-level $\hat{\beta}_i$ parameter in the immediate payment treatment (i.e. the constant in the OLS regression reported in Column 1 of Table 3) is equal to 1 (p value < 0.001). We cannot reject the hypothesis that the average individual-level $\hat{\beta}_i$ parameter in the end-of-day payment treatment is equal to 1 (p value 0.389).

[28] See, for example, Andreoni and Sprenger (2012a, b), Augenblick et al. (2015), Dean and Sautmann (2016), Halevy (2015), Epper (2015), Carvalho et al. (2016), Giné et al. (2017), Janssens et al. (2017).

but not individual preferences (Coller and Williams 1999). Second, even if one assumes that tradeoffs in experiments with money payments are driven by time preferences, should static preference reversals be interpreted as evidence of present bias? As discussed at length in Halevy (2015) and Giné et al. (2017), this interpretation assumes that subjects have stable utility functions and discount rates—that one's willingness to make tradeoffs between dated payments that arrive 1, 2, or 3 (or 100, 200, or 300) days from "now" does not depend on the calendar date upon which one is asked.

Providing definitive answers to these questions is beyond the scope of this paper. However, our experiment does allow for explicit tests of a number of the hypotheses under scrutiny. In what follows, we compare our results to those of other recent studies and present several additional pieces of analysis that speak to these questions. First, we test whether subjects behave in the manner predicted by standard models of intertemporal optimization, exploiting the high interest rates offered through the experiment to increase their present-discounted income stream. Second, we test whether subjects shift experimental payments toward dates when they expect to have limited liquidity.

5.1 Do Subjects Engage in Arbitrage?

Economists typically assume that people have access to perfect credit markets, and that their intertemporal tradeoffs are part of an optimal forward-looking consumption plan characterized by a set of Euler equations. However, most experimental economists—including those using lab experimental methods to study discounting—assume that subjects engage in narrow bracketing, viewing their dated lab experimental payments in isolation, as though they were consumption plans (Andersen et al. 2008; Rabin and Weizsacker 2009). Coller and Williams (1999) first highlighted the fact that experimental subjects with access to credit should engage in arbitrage, choosing immediate payments when the gross interest rate within the experiment is lower than their return on savings, and choosing delayed payments when the experimental return is lower than their cost of borrowing.

CTB experiments allow for explicit tests of the extent to which subjects engage in these types of arbitrage between lab and non-lab savings vehicles (Giné et al. 2017). Because many of the implicit interest rates offered in experiments are well above those available through the credit market, subjects who integrate lab payments into fully-optimal forward-looking consumption plans should only choose immediate consumption if they are liquidity-constrained. Those who hold liquid savings should allocate their entire endowment to the later payment date when the (implicit) lab interest rate exceeds the return on their savings (because this maximizes the net present value of their income stream).

We test this prediction in our data by dividing the sample into subjects known to have liquid savings in excess of the maximum early payment (the 28% of subjects who have more than 1000 Kenyan shillings, or 10.20 USD, in a savings account) and those who might be liquidity-constrained (those who do not have 1000 Kenyan shillings in a savings account). Figure 5 plots the proportion of subjects within each category who always allocate their entire endowment to the later payment date. The figure makes it clear that, though corner solutions are common, those subjects who are not liquidity-constrained do not appear to be engaging in arbitrage by consistently cashing in on the extremely high interest rates offered in the experiment. Across all interest rates, only 2.6% of subjects with liquid savings in the bank always allocate their entire endowment to the later payment date, versus 1.7% of those without 1000 shillings in the bank. Focusing on extremely high gross interest rates of at least 2 (i.e. guaranteed returns of at least 100% over a 2–4 week time horizon), which are very unlikely to be available outside the lab, we find that 13.8 percent of those who are not liquidity-constrained always allocate the entire endowment to the later payment date, versus 7.8% of those without 1000 shillings in the bank.[29] Thus, our data make it clear that most subjects are not engaging in the types of arbitrage that we would expect if lab payments were integrated into an optimal forward-looking consumption and savings plan.[30]

Next, we examine liquidity constraints. If our results were driven by liquidity constraints, we would expect our indicator of liquidity constraints to be correlated with the estimated individual-level $\hat{\beta}_i$ parameter. However, the individual-level regressions presented in Table 3 do not suggest that this is the case: the point estimate associated with the indicator for having at least 1000 Kenyan shillings in a liquid savings account is close to 0 and not statistically significant.[31] Moreover, it is not clear why liquidity constraints would have a differential impact on allocation decisions that depends on the timing of the experimental payments within the space of a few hours (i.e. whether payments arrive immediately after the experimental session and not at the end of the day). Thus, neither liquidity constraints nor arbitrage opportunities can explain our overall pattern of results.

5.2 Changes in Background Consumption

A more general mis-specification concern, suggested by Halevy (2015), is that consumption utility may not be stable over time. If background consumption reflects beliefs about future cash flow or liquidity, choice patterns that appear present-biased might result from subjects' attempts to shift payments toward dates when they anticipate that the marginal utility of consumption will be higher (for example, because they will have relatively little income or cash on hand), as suggested by Dean and

[29] Across all subjects, only 1.8% always allocate their entire endowment to the later payment date, and only 8.9% always do so at the higher interest rates.

[30] Interestingly, our results differ from those of Andreoni and Sprenger (2012a) in this regard: 70% of the CTB decisions observed in their experiment are corner solutions, suggesting that arbitrage may be a more reasonable explanation of behavior in that context. In our experiment, subjects who are not liquidity-constrained choose corner solutions 46.5% of the time; those who are potentially liquidity-constrained choose corner solutions 38.5% of the time.

[31] This finding is in line with earlier work by Meier and Sprenger (2010): in a sample of low-income tax-filers in the U.S., they find that "time preference measures are generally uncorrelated with credit constraints, future liquidity, or credit experience. This indicates that differential credit access, liquidity, and experience are unlikely to be drivers of experimental responses, and cannot explain the observed heterogeneity of present bias" (Meier and Sprenger 2010, pp. 202–203).

Sautmann (2016), Halevy (2015), and Ambrus et al. (2015). If subjects are shifting payments toward days when they expect the marginal utility of income to be high, static preferences reversals may or may not constitute evidence of dynamic inconsistency (Halevy 2015).

To test whether our results are driven by anticipated changes in the marginal utility of money payoffs, we added a series of questions about anticipated liquidity to the post-experiment survey. For each of the five possible payment dates in the experiment, we asked subjects whether they expected to have more cash-on-hand than normal, less than normal, or approximately the typical amount. These questions were only administered during the last 20 experimental sessions, so we have data from 324 of our 494 subjects.

In Online Appendix Table A8, we replicate the reduced form regressions reported in Table 1 in the sample of subjects who answered our liquidity questions, regressing the fraction of the endowment allocated to the earlier payment date on an indicator for having a front-end delay of 0 days, an indicator for the end-of-day treatment, an interaction between the two. As expected, results are similar to those in the full sample. In Columns 2 and 5, we add an indicator for having lower expected liquidity (cash-on-hand) at the later payment date, relative to the earlier payment date. If subjects were shifting their experimental payments toward periods of anticipated low liquidity, then the coefficient on our difference-in-liquidity variable should be negative and significant (because subjects will allocate less to the earlier payment date when they anticipate having lower background consumption on the later payment date). In fact, we find the opposite: the coefficient on our difference-in-liquidity variable is positive and significant. Moreover, including this variable has no impact on our main result: we continue to observe present bias in the immediate payment treatment (i.e. the amount allocated to the earlier payment date is higher when the front end delay is 0 days in the immediate payment treatment), and we do not observe present bias in the end-of-day payment treatment. Results are similar when we interact our difference-in-liquidity variable with the indicator for assignment to the end-of-day payment treatment (Online Appendix Table A8, Columns 3 and 6). Thus, we find no evidence that subjects reduce their allocation to the earlier payment date when they expect marginal utility to be higher at the later payment date, and we can rule out the possibility that such predictable changes in marginal utility explain our main results.

6 Conclusion

We conducted convex time budget experiments with truly immediate payment to test whether small delays of less than a day attenuate present bias. We exploited Kenya's mobile money technology M-Pesa to equalize transaction costs and uncertainty across all payment times and dates; we also introduced experimental variation in whether same-day payments were available immediately after the experiment or at the end of the day.

We find substantial evidence of present bias over money, but only when payments are truly immediate. Parameter estimates suggest that all delayed payments are discounted by between 7.6 and 9.8% in our immediate payment treatments. In sessions where the earliest possible payments arrive at the end of the day, we do not observe a statistically significant level of present bias (and we can formally reject the hypothesis that the degree of present bias is the same in both treatments). We also observe substantial heterogeneity in time preferences across subjects. However, variables associated with likely liquidity constraints (or the lack thereof) do not predict individual-level time preference parameters.

Our results point to a few broad conclusions about time preferencs. First, neither the standard model of intertemporal tradeoffs (in which lab payments are integrated into an optimal forward-looking consumption and savings plan) nor the simplest form of narrow bracketing (with background consumption equal to 0) provides an adequate description of the ways subjects make intertemporal tradeoffs in the lab. This suggests that researchers should devote greater attention to the development of experiments explicitly designed to test potential refinements of existing models, in order to build a more accurate positive description of the decision-making process.[32] Experiments intended to better understand the way humans use mental accounting and heuristics to simplify the process of organizing their financial lives seem likely to be particularly fruitful.

Second, our results suggest that (at least in some populations) present bias over money is important when the earliest payments are truly immediate; our findings are consistent with models such as the quasi-hyperbolic model (Laibson 1997; O'Donoghue and Rabin 1999) that predict a sharp distinction between immediate and slightly delayed payments. However, our experiment is not intended to trace out the relationship between individual discount factors and the level of the front-end delay over the minutes and hours after decisions are made. As such, we are unable to fully distinguish between the quasi-hyperbolic model and alternatives that do not suggest a clear discontinuity between immediate and nearly immediate payments (cf. Kirby 1997). We leave that to future work. It is also important to note that some other experiments involving immediate payments—most notably, Augenblick et al. (2015)—find little evidence of present bias over money in university student samples. The differences between these findings and our work suggest that some socio-economic groups may be more prone to arbitraging between lab and non-lab savings vehicles.

The development of the quasi-hyperbolic model of time preferences was motivated by the desire to provide a better positive explanation of the ways humans actually make intertemporal tradeoffs. However, since the advent of this tractable model of present bias, much of the energy has shifted toward exploring its predictions in field settings and testing between the quasi-hyperbolic model and the standard model. The new wave of theoretically-motivated time preference experiments, and the controversies generated by the conflicting results of these studies, highlights the need for more experimental work directed towards testing and refining the

[32] Ambrus et al. (2015) is an excellent example of work in this vein.

behavioral economic model of intertemporal decisionmaking, with a specific focus on not only the extent of present bias, but the ways subjects do or do not integrate controlled choices in the lab—or isolated decisions outside of the lab—into a larger, long-term optimization problem. Such an integrated framework has the potential to allow for a more complete reconciliation of existing results, and will also help policymakers seeking to nudge citizens and consumers toward achieving their long-term savings and consumption goals.

Acknowledgements We are grateful to Chaning Jang, James Vancel, and the staff of the Busara Center for Behavioral Economics for excellent research assistance, and to Ned Augenblick, Stefano DellaVigna, Pascaline Dupas, Ray Fisman, Jess Goldberg, Anett John, Shachar Kariv, Supreet Kaur, Maggie McConnell, Owen Ozier, Charles Sprenger, Dmitry Taubinsky, several anonymous referees, and numerous conference and seminar participants for helpful comments. This research was supported by Cogito Foundation Grant R-116/10 and NIH Grant R01AG039297 to Johannes Haushofer.

References

Abdellaoui, M., Kemel, E., Panin, A., & Vieider, F. M. (2018). *Time for tea now*.
Acland, D., & Levy, M. R. (2015). Naivete, projection bias, and habit formation in gym attendance. *Management Science*, *61*(1), 146–160.
Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International Economic Review*, *8*(1), 67–77.
Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496.
Ambrus, A., Asgeirsdottir, T. L., Noor, J., & Sándor, L. (2015). *Compensated discount functions: An experiment on the influence of expected income on time preference*. http://econ.duke.edu/people/ambrus/research.
Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583–618.
Andreoni, J., & Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753.
Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, *102*(7), 3333–56.
Andreoni, J., & Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, *102*(7), 3357–3376.
Augenblick, N., Niederle, M., & Sprenger, C. (2015). Working over time: Dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics*, *130*(3), 1067–1115.
Becker, G. (1962). Irrational behavior and economic theory. *Journal of Political Economy*, *70*(1), 1–13.
Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica: Journal of the Econometric Society*, *55*(3), 693–698.
Carvalho, L. S., Meier, S., & Wang, S. W. (2016). Poverty and economic decision-making: Evidence from changes in financial resources at payday. *American Economic Review*, *106*(2), 260–284.
Chakraborty, A., Calford, E. M., Fenig, G., & Halevy, Y. (2017). External and internal consistency of choices made in convex time budgets. *Experimental Economics*, *20*(3), 687–706.
Charness, G., Gneezy, U., & Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior and Organization*, *13*(Part A), 141–150.
Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007a). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, *97*(5), 1921–1938.
Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007b). Revealing preferences graphically: An old method gets a new toolkit. *American Economic Review*, *97*(2), 153–158.
Choi, S., Kariv, S., Müller, W., & Silverman, D. (2014). Who is (more) rational? *American Economic Review*, *104*(6), 1518–1550.

Clot, S., Stanton, C., & Willinger, M. (2015). *Are impatient farmers more risk averse? Evidence from a lab-in-the-field experiment in rural Uganda*. working paper.
Coller, M., & Williams, M. (1999). Eliciting individual discount rates. *Experimental Economics*, *2*(2), 107–127.
Dean, M., & Sautmann, A. (2016). *Credit constraints and the measurement of time preferences*. Working paper.
DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *American Economic Review*, *96*(3), 694–719.
Epper, T. (2015). *Income expectations, limited liquidity, and anomalies in intertemporal choice*. Universitat st. gallen discussion paper 2015-19.
Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692.
Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.
Fisher, I. (1930). *The theory of interest*. Augustus M. Kelley Publishers.
Fisman, R., Jakiela, P., & Kariv, S. (2017). Distributional preferences and political behavior. *Journal of Public Economics*, *155*, 1–10.
Fisman, R., Jakiela, P., Kariv, S., & Markovits, D. (2015). The distributional preferences of an elite. *Science*, *349*(5254), aab0096.
Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, *97*(5), 1858–1876.
Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.
Gabaix, X., & Laibson, D. (2017). *Myopia and discounting*. National bureau of economic research working paper no. 23254.
Giné, X., Goldberg, J., Silverman, D., & Yang, D. (2017). Revising commitments: Field evidence on the adjustment of prior choices. *Economic Journal*, *128*(608), 129–158.
Halevy, Y. (2008). Strotz meets allais: Diminishing impatience and the certainty effect. *American Economic Review*, *98*(3), 1145–1162.
Halevy, Y. (2014). *Some comments on the use of monetary and primary rewards in the measurement of time preferences*. Working paper.
Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica*, *83*(1), 335–352.
Harrison, G. W., Humphrey, S. J., & Verschoor, A. (2010). Choice under uncertainty: Evidence from ethiopia, india and uganda. *Economic Journal*, *120*(543), 80–104.
Harrison, G. W., Lau, M. I., & Rutström, E. E. (2013). *Identifying time preferences with experiments: Comment*. Center for the Economic Analysis of Risk Working Paper WP-2013-09.
Haushofer, J. (2014). *The cost of keeping track*. Working paper.
Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, *62*(6), 1291–1326.
Jakiela, P., & Ozier, O. (2016). Does africa need a rotten kin theorem? Experimental evidence from village economies. *Review of Economic Studies*, *83*(1), 231–268.
Janssens, W., Kramer, B., & Swart, L. (2017). Be patient when measuring hyperbolic discounting: Stationarity, time consistency and time invariance in a field experiment. *Journal of Development Economics*, *126*, 77–90.
Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, *126*(1), 54.
Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, *112*(2), 443–478.
Laibson, D., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., & Akerlof, G. A. (1998). Self-control and saving for retirement. *Quarterly Journal of Economics*, *1998*(1), 91–196.
Luhrmann, M., Serra-Garcia, M., & Winter, J. (2013). *Measuring teenagers' time preferences using convex time budgets*. working paper.
Meier, S., & Sprenger, C. (2010). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics*, *2*(1), 193–210.
O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103–124.
Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies*, *35*(2), 185–199.

Rabin, M., & Weizsacker, G. (2009). Narrow bracketing and dominated choices. *The American Economic Review*, *99*(4), 1508–1543.
Samuelson, P. A. (1937). A not on measurement of utility. *Review of Economic Studies*, *4*(2), 155–161.
Sun, C., & Potters, J. (2016). *Magnitude effect in intertemporal allocation tasks*. Working paper.
Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, *8*(3), 201–207.
Varian, H. R. (1982). The nonparametric approach to demand analysis. *Econometrica*, *50*(4), 945–972.
Varian, H. R. (1983). Non-parametric tests of consumer behaviour. *Review of Economic Studies*, *50*(1), 99–110.
von Bohm-Bawerk, E. (1890). *Capital and interest: A critical history of economical theory*. Macmillan and Company.
World Bank. (2015). *World development report 2015: Mind, society, and behavior*. International Bank for Reconstruction and Development.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.