World Development 159 (2022) 106029

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Can hypothetical measures of time preference predict actual and incentivised behaviour? Evidence from Senegal.

Jacopo Bonan [a], Philippe LeMay-Boucher [b,*], Douglas Scott [c]

[a] Department of Management, Economics and Industrial Organization, Politecnico di Milano, RFF-CMCC European Institute on Economics and the Environment (EIEE), Italy
[b] Department of Economics, Heriot-Watt University, Edinburgh, EH144AS, UK
[c] University of Oxford, Department of International Development, 3 Mansfield Road, Oxford, OX13TB, UK

ARTICLE INFO

Article history:
Accepted 6 July 2022

Keywords:
Time preferences
Discount rate
Randomized experiment
Senegal

ABSTRACT

Time preferences are an important determinant of decision-making and are widely measured through hypothetical survey questions. However, the extent to which they offer a good representation of time discounting remains largely unexplored. This paper estimates time preference parameters using a commonly-applied hypothetical elicitation method. We explore whether our estimated parameters correlate with actual and incentivized behaviours related to time preferences. First, we consider the correlation between our hypothetical measures and the result of an incentivised experiment using the unique reference numbers of banknotes as a means of determining an individual's willingness to save money. Individuals are given a banknote and informed that if they chose to retain this specific note for a randomly assigned period of time (2, 7 or 14 days) they will receive a second banknote, in effect doubling their initial endowment. Second, we consider the correlation between hypothetical measures and an individual's observable saving behaviour, including ownership of a savings account and participation in a Rotating Credit and Savings Association (ROSCA). Overall, our results show that hypothetically-derived time preference parameters are not significantly correlated with our measures of actual or incentivized behaviour. We explore the extent to which our results are due to limited power and find that a version of our results comparable to the relevant literature can detect effect sizes in line with similar studies. Furthermore, we recognise that our incentivized experiment will be a noisy reflection of time discounting and subject to confounding factors, such as the inherent fungibility of money. We provide ancillary evidence suggesting that our main results remain robust to these considerations and others.

1. Introduction

How individuals discount future rewards underlies the decisions they make in many aspects of their lives. If we look at studies with a focus on developing countries, time preferences have been notably linked to choices relating to savings behaviour and investment in agricultural inputs (see among others Ashraf et al., 2006; Duflo et al., 2011).[1] In the context of poverty alleviation, many have questioned the possibility that time preferences and self-control have a role to play in determining income and why some individuals remain poor (Lawrance, 1991; Atkeson & Ogaki, 1996; Harrison et al., 2002).

The importance of time preferences draws attention to the question of how these behavioural traits are measured. This study focuses on the use of choices over sets of hypothetical rewards, as a means of determining an individual's discount rate. It is difficult to evaluate precisely the extent to which hypothetical questions are used to elicit time preferences in studies relating to developing countries, however, the impression is that they remain widely used (D'Exelle, Van Campenhout & Lecoutere, 2012). A search of recent articles published after 2005, in development economics journals with international readership, yielded 27 articles that explicitly sought to measure time preferences. 17 of these used hypothetically-elicited discounting measures, 8 used real incentives (money or goods) and one used a combination of both (we could not determine which type of reward had been used in the remaining study). Choosing to use either hypothetical or real rewards represents an important experimental design choice. For example, in presenting respondents with hypothetical rewards, it is possible to offer a wider range of options and larger sums of

* Corresponding author.
E-mail address: P.LeMay-Boucher@hw.ac.uk (P. LeMay-Boucher).

[1] An already rich literature based on data from developed countries has found links between present-biased preferences and various economic decisions (see, among others, Meier & Spencer, 2010).

https://doi.org/10.1016/j.worlddev.2022.106029
0305-750X/

money, which allows for a deeper characterization of time preferences. This may not be feasible when using real incentives. However, with purely hypothetical rewards, respondents may have little incentive to work hard or thoughtfully on their responses (Harrison et al., 2005). Thus, the measures obtained can represent a noisy estimate of discounting behaviour.

Given these potential limitations, it is important to assess whether individuals' responses to hypothetical questions are a good predictor of actual behaviour. From the 17 papers we considered that employed hypothetical rewards, 13 of these investigated time preferences as independent variables. Overall, 5 of the 13 papers found a significant link between their proxy for time preferences and (at least one of) the dependent variables studied. For example, Ashraf et al. (2006) and Ashraf, Karlan and Yin (2010) find a significant correlation between time inconsistency and saving behaviours and attitudes.[2]

Our paper measures hypothetical time preferences via the standard Multiple Price list (MPL) format (Andersen et al. 2006) and first considers these results alongside an innovative experiment involving real money. We investigate whether hypothetical measures can be significant predictors of an incentivized decision expected to respond to time preferences. In order to do this, we develop an experiment, whereby individuals are given a banknote of 1,000 CFA francs (USD 2) (with the unique serial number recorded). Participants in the experiment are then informed that if they chose to retain this specific banknote for a randomly assigned period of time (2, 7 or 14 days) they will receive a second banknote, in effect doubling their initial endowment. We use the MPL questions to estimate a quasi-hyperbolic discounting specification (Benhabib, Bisin & Schotter, 2010) and generate a continuous measure of the discount rate and present-bias parameter, using a logistic function incorporating demographic variables (Tanaka et al., 2010). We then explore the extent to which these estimated time preference measures correlate with the probability of keeping the banknote, controlling for the duration of treatment (2, 7, or 14 days) and individual characteristics. We bootstrapped our standard errors, given the estimated nature of the time preference parameters. We find that an individual's hypothetical time preference choices are not significantly correlated with their choice to keep (or spend) the 1,000 CFA banknote. To further confirm this result, we also use our hypothetical measures to predict whether an individual owns at least one savings account (at a bank or microfinance institution) or participates in a ROSCA (Rotating credit and savings association).[3] Both are real-world outcomes that incorporate intertemporal choices and should be correlated with individual time preferences. Again, we find no significant correlation between the estimated parameters and the outcomes in question.

We explore the extent to which our (non-)results are due to issues related to power. In doing so, we use the estimated standard errors to construct ex-post Minimum Detectable Effects (MDEs) for our main results. Both as a test of robustness and to favour comparability with other studies, we also simplify the operationalization of our time preferences variables and the estimation technique. In place of continuous measures of an individual's discount rate and present bias, we employ binary measures to capture time preferences: i.e., patience (below the median estimated discount rate) and present-biasness (above the median estimated present bias parameter). In addition, while we maintain that the bootstrapped treatment of our errors is appropriate, we also present estimations using heteroskedasticity robust standard errors. In doing so, all correlations remain insignificant (at 5%), while MDEs lie between 17 and 24 percentage points, in line with the average MDEs computed from the directly comparable studies we found (approximately 24 percentage points). We interpret these findings as suggestive that our study is not unduly affected by low power, although we acknowledge that when we consider the estimated nature of our time preferences parameters, and bootstrap the standard errors, our MDEs become larger than the average of our comparable studies.[4]

We recognise that our experiment with incentivised behaviour is itself a noisy reflection of time preferences and subject to a number of confounding factors. For example, the inherent fungibility of money may have led participants, who would otherwise have spent the banknote, to substitute the equivalent amount of money from alternative household savings (or accessible savings/credit from friends, relatives or other contacts). Responses might also have been subject to a 'reputation effect', whereby individuals could have viewed the experiment as a test of their personal credibility and adjusted their behaviour accordingly. We discuss these issues at length below, where we argue that our main results remain robust to these considerations and others related to the definition of variables and specifications.

We also emphasise that this paper does not aim at directly comparing discount rates elicited with hypothetical questions and those with real rewards. The two exercises we employ are not directly equivalent and our banknote experiment was not designed with this intention. Our hypothetically-based MPL measures of discount rate and present bias are derived from a relatively large number of questions: up to 90 questions or 90 observation points per individual. Our banknote experiment only provides one observation point per individual and, as such, does not allow us to derive a measure of discount rate and present bias based on real rewards. To elicit any measure based on the banknote experiment we would have needed a (preferable large) number of observations, requiring us to conduct the experiment for each individual with different banknotes (perhaps adding 5,000 CFA and 10,000 CFA banknotes to the 1,000 CFA note) and using various periods of time between our visits. This was not planned nor attempted in the field. Given that our exercise is not one of comparison, we do not have a 'gold standard' in mind nor a preference towards an elicitation based on real incentives. Neither are we exploring the validity of different measures. We do not hold up the bank experiment as more informative than a hypothetically-elicited measure. Instead, we investigate whether a hypothetical measure, which is still widely used in the field, correlates with decisions based on real money. Intuitively, one would expect that it should.

To gain insight into the relationship between hypothetically-elicited measures of time preference and decisions related to tangible outcomes, it is useful to consider the small number of studies on time preference, involving both real and hypothetical rewards, most of which have been conducted under laboratory conditions: Johnson & Bickel (2002) and Madden et al. (2003) found no systematic differences in discount rates elicited using either type of

[2] In Appendix A we discuss the results from the literature review summarized in Table A1 to a greater extent.

[3] A ROSCA consists of a group of individuals who gather on a regular basis for a cycle of meetings. During meetings, members contribute a fixed amount of money to a common pot. This is then allocated to one member, who is then excluded from receiving the collective savings in subsequent meetings. However, he or she is still obliged to contribute to the pot for the remainder of the cycle. A cycle ends once every member has received the pot once. These groups are informal in the sense that they take place outside of the marketplace and are made without any binding legal arrangement.

[4] To alleviate any concern that our inability to reject the null of no correlation in our main results comes via relatively larger standard errors as a result of bootstrapping, we also provide estimates of our main results (using continuous measures of time preference parameters) with conventional Huber/White heteroskedasticity robust errors, essentially ignoring any variation in the estimated parameters. Our non-results are fully robust to this treatment (see Appendix Tables E3 and E4).

reward. Hinvest & Anderson (2010) found significantly higher levels of patience in participants offered real (versus hypothetical) rewards. Kirby & Maraković (1995) and Coller & Williams (1999) found a relatively lower discount rate for those responding to the hypothetical questions.[5] Outside of these laboratory studies, Ubfal (2016) provides another basis for comparison, through field-research conducted in Uganda. This paper concludes no significant variation between the two elicitation methods, suggesting minimal hypothetical bias. However, a review of the literature by Andersen et al. (2014) states that the evidence is overwhelming that there can be huge and systematic hypothetical bias when using this type of reward. Carvalho et al. (2016) seem to share this view by noting that it is clearly preferable to use incentivized tasks when possible.

The next section describes the context of our study, our elicitations methods and our experimental design. We then present an overview of the theoretical framework, our econometric models and discuss our results. A discussion of the measures taken to assess possible confounding factors follows with the concluding remarks.

2. Data and experimental design

2.1. Context of our study

Our study is based on a survey conducted in the city of Thies, Senegal, between May and July 2012. Thies is one of the largest cities in Senegal, with a population of approximately 240,000 inhabitants (at the time of the experiment). We use data collected on 360 randomly selected households across the whole territory covered by the city authorities. This represents an area of approximately 20 square km. We sampled the number of surveyed households, across all 24 Thies neighbourhoods, according to their respective share of the overall population estimates (based on the 2005 census). More information on our sampling methodology is available upon request.

For the purpose of this paper, a household is considered as a nuclear unit and consists of spouses, their children and all other members of the family who economically depend on the senior members. Our baseline survey was aimed at obtaining information on the general characteristics of each household member, including age, gender, level of education and ethnic affiliation. We also gathered information from the respondent concerning his/her work, monthly income, and a number of other characteristics (described in greater detail below). For 47% of the households surveyed the respondent was the household head, who is traditionally the husband.[6] In the remaining cases, the respondent was most often the spouse or (in a very few cases) another adult member of the household. We investigate the possible consequences of this below.

A brief overview of key variables obtained from the sample can be found in the first column of Table 1. To summarize, the majority of the respondents were female (63%) and 45 years of age, on average. The mean household income was around 210,700 francs CFA per month, which is equivalent to approximately US$443 (on the basis of the exchange rates at the time of the survey). Due to the sensitivity of obtaining income and salary levels, respondents were given a choice of 11 income intervals.[7] Therefore, income measures represent the mid-point in each interval, unless respondents provided more precise information. The three largest ethnic affiliations within the sample (Wolof, Poular and Sérer) approximately follow those of the country, as a whole.

2.2. Eliciting time preference parameters

Recent contributions to the time preference literature are often based on the Multiple Price list approach (MPL) proposed by Coller & Williams (1999). This method presents individuals with an ordered list of trade-offs between a fixed, immediate reward and an increasing future amount, subject to a specific period of delay. Given the relative simplicity of communicating this procedure to test-subjects, it is understandable that this approach is often favoured over more complex experimental designs. This is especially the case for fieldwork conducted in developing countries.

It has been suggested that measures obtained via this method may be susceptible to framing effects, dependent on the design of the price list employed (Harrison et al. 2005). We rely on multiple amounts and multiple time delays to mitigate these effects. All the questions used are of a yes/no type, allowing us to ask multiple questions to the same individual over the course of the interview. All respondents face the same protocol, and we assume that any biases in terms of understanding are not systematic across individuals.

The set of amounts and time delays used are shown in Table B1 of the appendix, all of which are purely hypothetical (with no real rewards attached). There are two possible values for the immediate reward: 10,000 CFA (approximately US$21) in panel A and 1,000 CFA (approximately US$2) in panel B. By way of comparison, we find that the mean of monthly income-per-capita for our sample of households is approximately 41,000 CFA (inclusive of members who are not economically active). Regarding the time-horizon, the set of choices start with a delay of 2 days, before increasing up to a period of 6 months (generating observations over periods of 2 days, 7 days, 14 days, 1 month and 6 months).

Following the standard methodology, the questions were designed to identify when the respondent switched from a (smaller) immediate reward toward a (larger) future reward. These questions were posed as follows: 'If you are sure to receive the sums mentioned at the given time, would you prefer accepting (X) francs CFA today or (Y) francs CFA in (t) days/months?' The delayed amount offered was then increased in subsequent questions until the respondent chose to switch. For example, in the case of the smaller initial reward (1,000 CFA, as opposed to 10,000 CFA), the first question proposed 1,000 CFA now and 1,000 CFA in two days. If the interviewee preferred the immediate reward (as would generally be the case), the delayed amount was increased to 1,050 CFA (US$2.19) and they were asked to express their preference again. This process was continued up until the point where the individual switched to the future reward. Beyond this point, we assume transitivity of preferences, such that the switching point is unique for any given initial amount (X) and time delay (t). Therefore, if an individual preferred a given amount in the future, compared to an initial value, he/she would also prefer larger amounts in the future (given the same time delay). In such an environment with low average education levels, a proper understanding of the time-discounting questions is particularly important. One round of pre-testing with 60 respondents was undertaken to limit any

[5] These studies consisted of a relatively small number of participants (<50), and crucially, actual and hypothetical rates were estimated for two separate groups of test subjects, thus making a direct comparison between both rates at the individual level more difficult.

[6] Different reasons can explain why only half of the household heads answered the questionnaire. In many cases, they did not live within the dwelling on a permanent basis, either visiting only for work-related reasons or to pay regular/irregular visits to the household. A limited number of heads did not have the time to answer the survey and delegated this responsibility to either their spouse or another adult. We did not meet a household that refused to take part in the survey.

[7] During our pilot, several individuals refused to give a precise value for their income, yet felt more inclined to answer if the question was presented as a choice of 11 multiple income brackets (from 0 to 250,000 CFA, in steps of 25,000 CFA, plus 1 choice of income greater than 250,000 CFA).

Table 1
Mean values of the main variables used in the analysis and results of an F-test for the equality of means.

	All		2-Day Treatment		7-Day Treatment		14-Day Treatment		F-test	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	p-values	
Household										
Household size	6.067	2.686	6.134	2.531	5.908	2.387	6.158	3.101	0.704	
Number of Children under 5	0.972	1.170	0.992	1.108	0.917	1.074	1.008	1.319	0.803	
Household income (100′000 CFA)	2.107	1.550	2.124	1.648	2.032	1.446	2.164	1.559	0.782	
Durables (# of items)[1]	7.933	4.656	7.546	4.424	7.533	4.231	8.717	5.194	0.103	
Respondent										
Gender (Male = 1)	0.370	0.484	0.370	0.485	0.400	0.492	0.342	0.476	0.647	
Age	44.883	13.637	44.664	12.011	45.583	14.214	44.400	14.614	0.794	
Respondent education (completed grades)	6.618	5.791	6.454	5.345	7.283	5.931	6.117	6.055	0.296	
In couple	0.866	0.341	0.849	0.360	0.892	0.312	0.858	0.350	0.571	
Owns home	0.755	0.431	0.756	0.431	0.733	0.444	0.775	0.419	0.756	
Ethnic Group										
Wolof	0.557	0.497	0.555	0.499	0.533	0.501	0.583	0.495	0.737	
Serer	0.106	0.308	0.118	0.324	0.092	0.290	0.108	0.312	0.799	
Poular	0.189	0.392	0.185	0.390	0.208	0.408	0.175	0.382	0.801	
Respondent is household head	0.474	0.500	0.504	0.502	0.508	0.502	0.408	0.494	0.212	
Estimated discount rate (r)	0.048	0.007	0.048	0.007	0.049	0.007	0.048	0.007	0.007	
Estimated present-bias parameter (β)	0.742	0.007	0.742	0.007	0.742	0.007	0.741	0.007	0.867	
Estimated risk-aversion (R)	0.545	0.167	0.525	0.146	0.536	0.163	0.575	0.185	0.061	*
Temptation[2]	0.262	0.440	0.210	0.409	0.208	0.408	0.367	0.484	0.010	**
Kept note[3]	0.783	0.413	0.874	0.333	0.800	0.402	0.675	0.470	0.001	***
Think will keep the note[4]	0.868	0.339	0.971	0.168	0.858	0.350	0.788	0.410	0.000	***
Savings account (bank or MFI)	0.591	0.492	0.487	0.502	0.617	0.488	0.667	0.473	0.015	**
Member of ROSCA	0.393	0.489	0.294	0.458	0.392	0.490	0.492	0.502	0.007	***
N	359		119		120		120			

Notes: The table reports an indication of the results of an F-test of the equality of means across treatment groups. *** p < 0.01, ** p < 0.05, * p < 0.1.
[1] This variable is the sum of a list of items owned by the household comprising: fridge, colour TV set, car, freezer, DVD player, sewing machine, gas cooker, stereo, bed (wood or metal), stove (camping stove), couch, clock, electric cooker, bicycle, gas lamp, oven, motorbike, petrol lamp, camera, charrette, electric fan.
[2] This binary variable takes value 1 if the respondent answered yes to Question 3: 'Do you think, yes or no, that you will have difficulties coping with the temptation to spend the banknote?' (0 otherwise). This variable is impacted by the number of days of the treatment (either 2, 7 or 14): the difference across groups is expected.
[3] This binary variable takes value 1 if the respondent did keep the actual note (0 otherwise). This variable is impacted by the number of days of the treatment (either 2, 7 or 14): the difference across groups is expected.
[4] This binary variable takes value 1 if the respondent answered yes to Question 1: 'Do you think that you can keep the money or not until the specified date?' (0 otherwise). This variable is impacted by the number of days of the treatment (either 2, 7 or 14): the difference across groups is expected.

systematic misunderstanding and the questionnaire was adjusted based on follow-up discussions.[8]

2.3. Eliciting risk preferences

Although this paper focuses primarily on the measurement of time preference, any non-instantaneous choices, from which an individual derives utility, are also likely to depend on levels of uncertainty regarding future outcomes (Andersen et al. 2008; Andreoni & Sprenger 2012). We thus follow Holt & Laury (2002) and administer another set of yes/no questions to elicit a measure of each individual's risk preference. Each respondent was offered the choice between two binary lotteries (A and B) involving gains (panel A) and losses (panel B, not shown), as outlined in Table B3 of the appendix (subjects always started from the top of each list). However, data obtained from panel B was scarce and therefore was not included in the calculation of risk preferences.[9] Lottery A is relatively riskier and has a higher payoff in the case of success. Lottery B is relatively safe and has a subsequently lower payoff in the event of a successful outcome. We set the probability of success the same for both the risky and safe lotteries and made the assumption of monotonic switching, in the sense that when an individual switched from lottery A to lottery B, as the probability of success decreased, he/she could not switch back to lottery A.[10] When eliciting risk preferences, we offered monetary payoffs based on a single task, selected at random from across the lotteries.

2.4. The banknote experiment

Following the baseline questionnaire, each respondent was given a 1,000 franc CFA banknote.[11] The unique reference number of this note was recorded, and the individual was informed that, if they produced the same banknote when the household was visited on a second occasion, they would receive another 1,000 francs, and could retain both notes. The specific date of the second visit was randomly assigned, as 2 days, 7 days or 14 days from the initial visit, and this was announced to each household.[12] One household in three was assigned to each of these three possible treatment groups.

[8] Panel A of Table B2 (Appendix B) shows the extent of violations of the transitivity assumption during the pilot survey, i.e. when an individual switched more than once between the present and the future rewards. Only one of the 60 individuals displayed such inconsistent behaviour.

[9] This was due to a significant fraction of individuals showing reluctance in providing answers to this part of the experiment (as it involved losses), even when we repeatedly explained that the experiment was purely hypothetical.

[10] Panel B of Table B2 (Appendix B) shows violations of the monotonic preference assumption for the lottery task in the pilot study, i.e. when individuals switched more than once between the two lotteries. Only three individuals displayed this behaviour in the pilot.

[11] The order of the two tasks (MPL questions and the banknote experiment) was not randomized.

[12] Because of organizational, time and resource constraints, it was not feasible for the same individual to play more than one scenario from the MPL. In particular, it was not possible to offer the same individual a banknote of 1,000 CFA and a banknote of 10,000 CFA (or offer second visits over different periods of time). It should also be apparent that any experiments conducted with actual rewards will clearly be limited by financial constraints. For example, were the experiment conducted using a 10,000 CFA banknote, given the proportion of the sample who retained the note in the 1,000 CFA case (and noting that this retention rate could be higher for the larger payments), the basic costs of conducting this experiment alone would, unfortunately, have been beyond our means.

However, one individual refused to participate in the experiment, reducing the overall sample size to 359. Table 1 shows a comparison of the mean values for key variables within the three treatment groups (2 days, 7 days and 14 days). From our 11 potential baseline controls (shown in the upper section of Table 1), no significant differences are observed across the groups. However, there exists some heterogeneity between treatments in our estimated measure of risk aversion and whether or not a household member owns a savings account or is a member of a ROSCA (Rotating credit and savings association).[13] All other variables which differ significantly between treatments are functions of the treatment duration and are, therefore, expected to differ between assignments.

Perhaps, a more intuitive treatment would have been to offer each respondent *either* a 1,000 franc CFA note today or 2,000 francs CFA in *t* days during a second visit. This would have represented a replication of the MPL questions. However, during our piloting phase, this approach proved difficult to implement cleanly. When presented with this choice, some respondents opted for the immediate reward because they perceived that there was a possibility that we would not return for the second visit. This was the case in all three treatments (2, 7 and 14 days), in spite of our efforts to assure the participants that our second visit would take place. Our pilot survey indicated that our results were likely to be biased by this possible uncertainty over future rewards, were we to attempt to implement this approach. We found that by initially offering a note of 1,000 CFA surveyed individuals were not inclined to think that our second visit was in any doubt, even in the 14-day treatment group (who would experience the longest period between visits). Offering money during our first visit gave credibility to our experiment, such that this present treatment approach allowed us to avoid 'trust' bias, whilst also allowing for easier implementation. We discuss the issue of trust in the next subsection.

We are aware that our banknote experiment not only provides an indication of time preference but also will be tainted by how individuals cope with temptation when saving money for short periods of time (how good they are at committing). These effects are difficult to disentangle, but in an attempt to do so, we use additional information obtained alongside our time preference parameters (see below). It is also important to re-emphasize that our experiment is not a replication of our hypothetical MPL questions. As such, we are not testing the validity of hypothetical versus incentivized time preference measures. Our goal is rather to check if our hypothetical time preference parameters correlate with incentivised behaviour.

Once the banknote was received, each individual was asked a series of five questions:

Question 1: 'Do you think that you can keep the money or not until the specified date?'

Question 2: 'Why do you think you can or cannot?'

Question 3: 'Do you think, yes or no, that you will have difficulties coping with the temptation to spend the banknote?'

Question 4: 'Do you plan, yes or no, to do something in order to make sure that you will not spend the note?'

Question 5: 'If yes what?'

The first three questions were aimed at determining to what extent an individual believed they could resist temptation (and deal with self-control issues) during the experiment (the third question addressed this specifically). The last two questions were intended to identify any mechanisms they planned to use to ensure they avoided temptation and allow us to check if respondents were considering using any form of commitment device to ensure they did not spend the money. These last questions were also designed to evaluate any potential bias in behaviour, due to the inherent fungibility of the reward. That is to say, we wanted to see how likely the participants were to consider drawing money from an existing pool of cash (or credit), in order to increase their expenditure now, while still managing to retain the specific banknote provided. Were this form of expenditure-source switching common within the experiment, the results obtained could be misleading. Our descriptive statistics show that only 1.5% (3 out of 205) of our respondents, who answered 'yes' to *Question 4*, would consider using such liquidity (or borrowing), in order to help them to keep the specific banknote. Answers indicative of this were: 'I will borrow around me (from friends or acquaintances) if I need, instead of using the note.' None mentioned the use of savings in ROSCAs, microfinance institutions (MFIs) or bank accounts. With such a small sum (1,000 CFA), we argue that this reasoning is likely to be marginal. However, the issues of fungibility and self-control are discussed at greater length in Section 5.2.

Table 1 shows that those who retained the note (variable 'kept note') accounted for 78% of the overall sample. This proportion declines from 87% for a delay of 2 days, to 80% (for 7 days) and 67% for 14 days. These differences are significant between 2 and 14 days, and 7 and 14 days, but not between the 2- and 7-day treatments. 87% of our sample indicated that they thought they could keep the note until the specified date (yes to *Question 1*; variable 'think will keep the note'). As would be expected, this proportion is diminishing significantly with the number of days in the treatment (97% for the 2-day treatment, 86% for the 7-day treatment and 79% for the 14-day treatment). Answers to *Question 3* indicate that 26% of all respondents think that they will experience difficulties coping with the temptation to spend the banknote (variable 'Temptation'). This proportion is significantly larger for those within the 14-day treatment (37%) than for either the 2- or 7-day treatments (21%).

A large majority of the answers to *Question 2* (following a positive answer to *Question 1*) highlighted the importance of gaining an additional 1,000 CFA as the primary motivation for keeping the banknote. Answers to *Question 2* (from those who believed themselves *unable* to keep the banknote) mostly indicated that debts needed to be repaid or that urgent familial needs would prevent them from saving the note. 57% (205 out of 359) of the respondents indicated in *Question 4* that they planned to do something in order to make sure that they would not spend the banknote. Of this 57%, answers to *Question 5* indicated that 23% (48 out of 205) intended to give the note to somebody they trusted, in order to prevent them from using it, and 33% (68 out of 205) intended to hide the note somewhere safe (under their mattress, cupboard, etc.). In general, the pattern of these responses is indicative of the findings of Ashraf, Karlan & Yin (2006), and Dupas & Robinson (2013a), who underline the importance of simply having access to a safe place to keep money, as a means of increasing savings.

2.5. The issue of trust

The issue of 'trust' applies specifically to the second visit. It did not take the form of mistrust or suspicion towards our project more generally, or our team of well-trained and experienced local enumerators. Out of the 360 households we approached, none refused to be surveyed, and enumerators felt that they experienced a similar welcome as they had in previous surveys in the same region. The enumerators we worked with were senior and experienced enough to establish trust and respect from surveyed individ-

[13] The reason why we observe these differences, given our experimental design, is most likely due to the relatively small sample sizes. There was no differential refusal rate to participate in the study by treatments. As far as we can tell, none of our enumerators showed strategic behaviour in selecting households, and our assignment of treatments was conducted in a proper way that should have prevented this outcome.

uals and knew how to approach local people, as they had previous experience conducting surveys in the city of Thies. This respect and trust helped mitigate any remaining concerns over missing monetary incentives.

Given the design of our banknote experiment, we could not implement a simple 'front-end delay' method to control for potential confounds due to lower credibility associated with future payments (see Harrison, Lau & Williams, 2002). Specifically, if participants are in any doubt that they will receive a reward in the future, they may prefer a current reward, irrespective of their actual discount rate. As our experiment was based on an early gift of a banknote of 1,000 CFA, a 'front-end delay' approach was not applicable. Nevertheless, during our pilot survey, we experimented with a small subsample of around 30 individuals, to determine if their behaviour was different if we visited them the day after we gave them the banknote (with an unannounced visit). We did this for all three randomly assigned delays (2, 7 and 14 days). It was designed to improve our credibility and to signal the seriousness of our next planned visit. Our anecdotal results show that, overall, individuals in both subsamples reacted similarly whether they received this impromptu visit or not. Given this, and our limited resources, we minimized survey costs by not generalising this visit to our large sample of 360 individuals.

3. Empirical analysis

Our empirical analysis has two components: 1) the estimation of time preference parameters based on the MPL questions; 2) establishing whether these parameters are correlated with the choice made in the banknote experiment or participating in other saving activities (ownership of a saving account or taking part in a ROSCA).

3.1. Estimation of the time preference parameters

Various specifications have been considered which allow for relative impatience over short-term rewards. Many of these models are based around 'hyperbolic' or 'quasi-hyperbolic' functional forms (see Laibson, 1997) and have often been found to fit the data more accurately than standard, exponential discounting.[14] Benhabib, Bisin & Schotter (2010) provide a general expression for an individual's discount factor, which allows for testing among possible models, namely exponential, hyperbolic and quasi-hyperbolic discounting. We use this nested formulation as a starting point for our estimation of time preferences.

$$D(y,t,\beta,r,\theta = \begin{cases} 1 if t = 0 \\ \beta(1-(1-\theta)rt)^{\frac{1}{1-\theta}} if t > 0 \end{cases} \quad (1)$$

In equation (1), the discount factor *D(y, t)* is the value that makes an individual indifferent between two alternative time/reward pairs *(y D(y, t), 0)* and *(y, t).* The discounted utility model involves a linear utility function. In addition to the time between rewards t and the underlying discount rate r, this discount factor is expressed as a function of the parameters β and θ, which are intended to characterise the various forms of discount function considered within this study. Specifically, β is a parameter representing present bias (in a quasi-hyperbolic specification) and θ parameterizes the curvature of the discount function. Dependent on the restrictions imposed on the parameters, this specification can represent various forms of time preference, through nesting exponential, hyperbolic and quasi-hyperbolic discounting functions, as follows.

i) When β = 1 and θ is approaching 1, equation (1) represents exponential discounting (e^{-rt}), whereby the discount factor decreases over time at a constant rate.
ii) When β = 1 and θ = 2, equation (1) represents pure hyperbolic discounting ($1/(1 + rt)$). In this case, the discount rate decreases over time and displays a non-constant absolute rate of change.
iii) When θ is approaching 1, equation (1) displays future rewards under quasi-hyperbolic discounting (Laibson, 1997). $D(y, t, \beta, r, \theta)$ takes on the form βe^{-rt}, allowing for an individual to display a 'present bias' towards immediate reward, with all non-immediate amounts discounted by a factor β.

In our early attempts to estimate the most general form of the discounting equation described above (with an unrestricted θ), we found little improvement in the explanatory power of our model, using this relatively more complex function. As a result, we opted to employ the quasi-hyperbolic discounting specification, with two unrestricted parameters r and β. A comparison of the estimated discount parameters between four alternative discounting equations is presented in Table C1 of the appendix.[15] Following Tanaka et al. (2010), the term μ is included as a response-sensitivity (noise) parameter.

$$P(X > (Y,t)) = \frac{1}{1 + e^{\mu(X - Y\beta e^{-rt})}} \quad (2)$$

Under the assumption of quasi-hyperbolic discounting, equation (2) shows the probability that the immediate reward X (at time 0) is preferred to the delayed reward Y (at time t). The parameters of the values of t, X and Y are obtained from the time delays and amounts proposed in the various MPL questions. Quasi-hyperbolic discount parameters r and β are estimated within the model, using the logistic function shown in (2). These parameters, will depend on a vector of characteristics, assumed to be correlated with individual time preference.

3.2. The correlation with incentivized and observable behaviour

In the second part of our analysis, we initially estimate a set of regressions to establish the correlation between our estimated time preference parameters and the actual choices made by individuals during the banknote experiment. Within these regressions, the dependent variable takes the value 1 if the individual kept the banknote (and waited for the next visit to receive the second payment), and zero if the individual could not produce the banknote at the later time. The motivation behind these regressions is to investigate the role played by the estimated time preference parameters in determining incentivised behaviour. Again, we re-emphasize that we are not comparing the validity of hypothetical and real rewards. Our hypothetically-based MPL measures of discount rate and present bias are derived from a relatively large number of questions: up to 90 questions (90 observations) per individual. Our banknote experiment only gives one observation point per respondent and, as such, does not allow us to fully characterize an individual's time preferences. Our objective is only to determine if hypothetically-elicited time preference parameters are correlated with the decisions individuals make in our incentivized banknote experiment.

Having determined the correlation between our time preference measures and the outcome of the banknote experiment, we

[14] Some notable examples of studies that reject the exponential discounting form include Rachlin et al. (1991), Kirby & Maraković (1995) and Myerson & Green (1995).

[15] In this regard, our analysis is similar to Tanaka et al (2010). They estimate the unrestricted version of equation (1) but find that it adds little to the explanatory power of the model (compared to the estimation of the quasi-hyperbolic specification), and so focus attention only on the quasi-hyperbolic discounting version of equation (1).

further assess the predictive strength of our hypothetically-elicited measures by assessing their ability to predict whether an individual owns a saving account (at a bank or microfinance institution) or participates in a ROSCA.

4. Results

4.1. Time preference parameters

Panel A of Table 2 indicates the proportion of respondents who switched at the corresponding future amounts (in the respective time period) in the lower (1,000 CFA) initial amount MPL questions. For the 6-month time frame, almost all individuals (96%) preferred the immediate reward to all amounts offered (ranging from 1,050 to 3,000 CFA). This 'no switch' proportion reduces to 84%, when the time delay is reduced to 1 month, and decreases further as the delay approaches the present. Table 2 also indicates that 88% of the sample preferred at least twice the initial amount when the time delay was 7 days, while 92% preferred at least 1.5 times the initial amount over the shortest time period stipulated (2 days). These hypothetical results appear to show a high degree of impatience. For example, approximately 63% of those sampled were unwilling to accept any of the given future rewards in 14 days, even when the opportunity of tripling their initial endowment was proposed.

The results from the MPL questions for the higher (10,000 CFA) initial amounts, shown in Panel B, also indicate a degree of impatience in the hypothetical questions. For 86% of the sample, none of the future rewards offered were sufficient to induce them to accept a delay of 6 months. This percentage falls substantially, however, as the time delay for future rewards is reduced, such that only around 8% of individuals were unwilling to forgo the immediate payoff over the 2-day time horizon. A comparison of results, between the high and low initial rewards, indicates that the proportional increase required to induce a switch to the future payoff is generally lower in the 10,000 CFA questions (in all time frames). This is evidence of a commonly observed 'magnitude effect', whereby small amounts are discounted more heavily than relatively larger rewards (Kirby & Maraković, 1995; Coller & Williams, 1999).

For each initial amount in the MPL questions (1,000 and 10,000 CFA), 5 time periods were considered (2 days, 7 days, 14 days, 1 month and 6 months), with 9 questions per time period (corresponding to 9 delayed reward values). Therefore, for each of the participants, MPL questions provided a selection of 45 responses for each of the two initial amounts of money proposed, giving us a total of 90 observations per individual and 32,310 observations overall from the sample of 359 respondents. Due to our assumptions on preference transitivity, an individual could switch between immediate and future rewards at most once in each time period (providing only one point of variation for each of the given X,t combinations).

We estimate our discounting parameters in equation (2) following Tanaka et al. (2010) and incorporate demographic variables in the logistic function directly. This requires defining β and r in equation (2) as a function of individual and household characteristics: $\beta = \alpha_0 + \Sigma\, \alpha_i X_i$ and $r = \gamma_0 + \Sigma\, \gamma_i X_i$, where X_i represents these characteristics and α_i and γ_i their associated coefficients. As the time preference parameters are estimated from a sample of 90 observations per respondent and the two initial amounts (1,000 and 10,000 CFA) generate a clear source of heteroskedasticity, cluster robust standard errors at the individual level are reported in Table 3.

Our results indicate that, within our sample, only the respondent's gender and household size are significant determining factors in predicting the parameters β and r (at $p < 0.05$), with a male respondent being associated with a relatively higher discount rate and a lower degree of present-bias (the parameter β is negatively related to the degree to which individuals favour immediate reward). There is also evidence of a (minimal) positive effect of an additional household member (variable 'household size') on an individual's discount rate.

The estimated coefficients in Table 3 are used to compute fitted values of the parameters β and r, relating to each of the individuals. Table 4 shows a summary of the sample means for these predicted variables, with the mean estimated values of the underlying discount rate r at 4.8% and the present bias parameter β at 0.74. As the t parameter in model (2) is measured in number of days, our estimations of r represent daily discount rates. This implies that the average individual in the sample should be indifferent between 1,000 CFA today and either 1,488 CFA received in 2 days; 1,891 received in 7 days or 2,646 CFA received in 14 days.

4.2. Comparing responses to the MPL and the banknote experiment

Before comparing results from the banknote experiment and MPL questions, it is important to re-emphasise that both our hypothetical and incentivised tasks are capturing different decision processes. With this caveat, we offer here a simple comparison between the answers provided to the MPL questions and the corresponding behaviour observed in the banknote experiment. For each participant, one question within the MPL replicates the exact time frame and reward pair offered within the banknote experiment.[16] For example, an individual in the 2-day treatment, who claimed in the MPL questions to prefer 2,000 CFA in two days to 1,000 CFA now, should reasonably be expected to keep the banknote and receive the additional 1,000 CFA two days later (if behaviour is consistent).

Table 5 shows whether the corresponding MPL question was able to predict an individual's behaviour in the banknote experiment using this simple comparison. Our results indicate that the ability of the MPL to predict behaviour declines as the time-horizon increases. The consistency between the MPL and the banknote experiment is 69% for the 2-day treatment group but decreases to only 38% for the 14-day treatment. Table 5 indicates that a substantial share of the sample (44%) retained the banknote, having given responses to the MPL questions indicating that they would need a larger remuneration than 1,000 CFA (offered in our experiment) to wait for the stipulated period (2, 7 or 14 days). The proportion of such individuals also increases with the time delay, from 24% in the 2-day treatment to 59% in the 14-day treatment. It is this form of inconsistency (impatience in the MPL, but patience in the banknote experiment) that dominates the results in Table 5. Again, however, we stress that the two decisions differ by more than the type of reward offered.

4.3. Risk aversion parameter

Following the discussion in section 2.3 (regarding the influence of uncertainty in choices over future outcomes), our measure of risk aversion is elicited using the lotteries described in Table B3 of the Appendix. Respondents were asked to choose between a relatively risky (choice A) and a relatively safe lottery (choice B), with the probability of success communicated to individuals using a bag containing different combinations of two colours of marbles (from which one marble would be drawn to determine the outcome). As

[16] This would be from the selection of questions where X = 1,000 CFA and Y = 2,000 CFA. One question within this set would ask the subject to choose between these two amounts, over a period of time matching their treatment assignment in the banknote experiment.

Table 2
The proportion of respondents who opted for the future reward at the indicated amount.

Amount	2 days (%)	7 days (%)	14 days (%)	1 month (%)	6 months (%)
Panel A: With an initial option of 1,000 CFA (≈US$2); N = 359					
1,000	0.6	0.3	0.3	0.3	0.3
1,050	0.3	0	0	0	0
1,100	1.4	0.6	0	0	0
1,250	5.8	1.4	0	0	0
1,500	31.2	5.0	1.9	1.1	0.6
1,750	8.4	5.0	1.4	0.3	0.3
2,000	22.3	19.5	8.1	1.9	0.6
2,500	7.0	15.9	10.3	2.2	0.3
3,000	6.1	19.2	15.3	10.3	1.7
'No switch'	17.0	33.1	62.7	83.8	96.4
Panel B: With an initial option of 10,000 CFA (≈US$20); N = 359					
Amount	2 days (%)	7 days (%)	14 days (%)	1 month (%)	6 months (%)
10,000	1.4	0.3	0.3	0.3	0.3
10,500	3.9	1.1	0.6	0.3	0
11,000	10.9	3.1	0.8	0.3	0.3
12,500	25.1	7.0	3.9	1.4	0.6
15,000	30.6	20.9	10.6	3.9	1.1
17,500	7.2	15.3	6.4	2.5	0
20,000	9.5	25.9	21.2	10.6	1.4
25,000	1.4	8.1	12.3	9.2	1.4
30,000	1.7	8.1	18.9	15.9	8.9
'No switch'	8.4	10.3	25.1	55.7	86.1

Table 3
Estimated time preference parameters.

	r (1)	β (2)
Gender (Male = 1)	0.0135***	0.0128***
	(0.0005)	(0.0043)
Age	−0.0000	−0.0002
	(0.0000)	(0.0002)
Household size	0.0002**	0.0007
	(0.0001)	(0.0008)
Number of Children under 5	0.0002	−0.0022
	(0.0002)	(0.0019)
Income (in 100,000 CFA)	0.0002	−0.0005
	(0.0002)	(0.0014)
Durables (sum of items)	−0.0001	−0.0000
	(0.0000)	(0.0005)
Respondent education (completed grades)	0.0000	0.0001
	(0.0000)	(0.0004)
μ	−5.4070***	−5.4070***
	(0.1453)	(0.1453)
Constant	0.0422***	0.7431***
	(0.0008)	(0.0080)
Observations	32,310	32,310
Respondents	359	359
R^2	0.474	0.474

Notes: Individual-level, cluster robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.
These results are based on all 359 respondents for all available price list framing: for both 1,000 CFA and 10,000 CFA for 2, 7, 14, 31 and 186 days.

would be expected, the fraction of individuals choosing the risky lottery (A) declines as the probability of the higher payoff decreases. This reflects, in part, the change in the expected income difference between the risky and safe lotteries, which falls from 480 CFA to 180 CFA, as the probability of the higher payoff falls from 0.8 to 0.3. A rational, expected-utility maximizing individual, with weakly risk-averse preferences, should switch from choosing the risky to the safe lottery at most once over the course of the six tasks.

We make the assumption that an individual's preferences over outcomes in this lottery can be represented by a constant relative risk aversion (CRRA) utility function of the form $u(x) = \frac{x^{1-R}}{1-R}$. This function is used to place bounds on the CRRA coefficient R. An individual choosing the risky lottery in all tasks must have $R \leq 0.22$, whereas choosing the safe lottery in all tasks implies $R \geq 0.82$. Those who switch from the risky to the safe lottery between tasks 1 and 6 will have a value of R bounded within a strict subset of the interval (0.22, 0.82). Based on a respondent's choices in the lottery, the risk aversion parameter R is calculated directly from the CRRA function above. Table 6 shows these values and their frequencies within the sample. However, 7% of individuals did not switch at any point in the experiment; the risk aversion parameter for these participants is set at 0.22. Table 6 indicates a mean estimate of R at 0.55, with a standard deviation of 0.17, values which are broadly in line with results in Harrison et al. (2010).

4.4. Correlation between hypothetical measures and the outcome of the banknote experiment

In the second part of our analysis, we determine whether each individual's estimated values of β and r (present bias and discount rate) predict the behaviour we observe in the banknote experiment. We use these variables as regressors in a set of estimations, where the dependent variable represents whether or not the banknote was kept (1 if the note was shown to our enumerator upon

Table 4
Summary of the means of the time preference parameters.

	Individuals	Mean	Standard deviation	Min	Max
Discount rate (r)	359	0.048	0.007	0.042	0.060
Beta (β)	359	0.742	0.006	0.728	0.759

Notes: These results are based on all 359 respondents, for all available price list framing: for both 1,000 CFA and 10,000 CFA, and for 2, 7, 14, 31 and 186 days.

Table 5
Consistency of the MPL preference questions and behaviour.

	All		2-day		7-day		14-day	
	# Ind	%	# Ind	%	# Ind	%	# Ind	%
MPL predicted banknote experiment	179	50	82	68.9	51	42.5	46	38.3
Note saved (predicted) [1]	122	34	76	63.9	36	30	10	8.3
Note spent (predicted) [2]	57	15.9	6	5	15	12.5	36	30
MPL did not predict banknote experiment	180	50	37	30.1	69	57.5	74	61.7
Note saved (not predicted) [3]	159	44	28	23.5	60	50	71	59.2
Note spent (not predicted) [4]	21	5.8	9	7.6	9	7.5	3	2.5
Total	359	100	119	100	120	100	120	100

Notes: Figures represent the results of a comparison between an individual's behaviour in the banknote experiment and their response to the specific question in the MPL corresponding to the equivalent time delay and payments that they faced in the incentivized experiment (1,000 CFA now, 2,000 CFA at time *t*).
[1] In the MPL, 2,000 CFA at time *t* is preferred to 1,000 CFA now, and the banknote was kept for *t* days.
[2] In the MPL, 1,000 CFA now is preferred to 2,000 CFA at time *t*, and the banknote was not kept for *t* days.
[3] In the MPL, 1,000 CFA now is preferred to 2,000 CFA at time *t*, but the banknote was kept for *t* days.
[4] In the MPL, 2,000 CFA at time *t* is preferred to 1,000 CFA now, but the banknote was not kept for t days.

Table 6
The distribution of the estimated risk aversion parameter (R).

Value for R	Frequency	Percentage
0.22	25	6.96
0.30	23	6.41
0.44	99	27.58
0.56	99	27.58
0.67	55	15.32
0.77	13	3.62
0.82	45	12.53
Total Observations	359	100

the second visit, and 0 otherwise). The results described below are obtained via probit regressions, yet our results are similar when using logit or ordinary least squares (OLS) estimations.

Table 7 reports the correlation between the estimated discount rate and present bias (from the hypothetical MPL questions) and the observed results of the banknote experiment. The estimated values of r and β are standardized in all estimations in Table 7, such that the coefficients indicate the effect of a 1 standard deviation change in the estimated discounting parameters. Before introducing the time preference parameters, model 1 initially reports only the effects of the treatment durations (2, 7 and 14 days) on the outcome in the experiment. In model 2, estimated individual discounting parameters (discount rate and present bias) are added as explanatory variables. Model 3 also controls for differences in risk aversion, through the inclusion of the estimated variable R, while model 4 reports the effects of our hypothetical measures without controlling for the time delays. Model 5 introduces additional controls, which could be expected to influence the outcome of the experiment (other than through time preference).[17] The final model also includes an indicator intended to capture self-control, based on the question 'Do you think, yes or no, that you will have difficulties coping with the temptation to spend the banknote?' We interact this binary variable with treatment duration, as this question was posed after we assigned our respondents to the 2-, 7- or 14-day treatment groups (implying their response will be a function of treatment assignment). In all models, we include a neighbourhood-level fixed effect to mitigate the influence of unobserved, time-invariant characteristics relating to the 24 neighbourhoods in Thies.

Standard errors in Table 7 are bootstrapped, due to the presence of the generated regressors β, r and the risk aversion variable R. As the time preference parameters represent predicted values from the first part of the analysis, any heteroskedasticity in the first stage errors, as a consequence of the use of two initial MPL amounts (1,000 and 10,000 CFA), will not invalidate our inference. All columns in Table 7 report marginal effects (at the mean), with associated standard errors in parentheses. Terms reported in the square brackets indicate Minimum Detectable Effects (MDEs).

As would be expected, across all models, the 14-day treatment frame consistently shows a strong, negative correlation with the probability of the banknote being kept (relative to the 2-day treatment base category). Inclusion in the 14-day treatment reduces the probability of the note being retained by between 19 and 22 percentage points in all regressions (relative to the 2-day group). The 7-day frame displays the expected sign, a smaller coefficient, but is not statistically significant compared to the 2-day group. We also find the probability of retaining the note is significantly different between the 7- and 14-day treatments (p-value = 0.045, based on column 1 of Table 7). These results are in line with the proportions of individuals retaining the note over the different treatment periods (see Table 1).

The most striking result in Table 7, however, is that the estimated effect of the discount rate and present bias variables r and β is never significant (in any of our models). The same result applies when we test for the joint significance of these coefficients. The tests of the joint hypothesis, (reported at the bottom of Table 7) generate p-values well in excess of 10% in all estimations. Among the additional variables added in column 5, few seem to be significantly related to the probability that the banknote would be retained, while our measure of risk aversion also appears non-significant in most estimations in Table 7. The indicator variable for temptation (see section 2.4) in column 6 does appear to influence the outcome of the experiment, however, with those reporting that they expect to experience difficulty avoiding temptation more likely to spend the banknote in the 2- and 7-day treatment durations. The effect of beliefs about temptation appears less important (and less significant) over the longest treatment period (although, the coefficient does display the expected sign).[18]

It is worth mentioning that our computations at the design phase show that our tests for the coefficients of our two treatments (14 days and 7 days) could detect expected effect sizes of 16 percentage points, with power above the widely considered satisfac-

[17] In selecting these controls, much care was taken to avoid a high degree of multicollinearity with the generated time preference parameters. Were we to include too many of the regressors from the first stage of our estimations, co-movement between these variables and the time preference measures would lead to imprecision in the second stage coefficients, potentially undermining our ability to detect a significant effect.

[18] For each of our treatments, we have the following number of individuals who believe they will experience temptation: 2-day (25 out of 119); 7-day (25–120); 14-day (44–120).

Table 7
The estimated effects of time preference parameters on the probability of keeping the banknote (probit regressions, marginal effects reported).

	(1)	(2)	(3)	(4)	(5)	(6)
	The dependent variable is Banknote kept = 1 (=0 otherwise)					
Discount rate (r)		0.061 (0.122) [0.34]	0.064 (0.123) [0.34]	0.058 (0.114) [0.32]	0.009 (0.160) [0.45]	0.002 (0.171) [0.48]
Present bias parameter (β)		−0.053 (0.122) [0.34]	−0.058 (0.123) [0.34]	−0.052 (0.113) [0.32]	−0.048 (0.160) [0.45]	−0.051 (0.171) [0.48]
Risk aversion (R)			−0.184 (0.160)	−0.267* (0.162)	−0.141 (0.163)	−0.170 (0.160)
Income (in 100′000CFA)					0.011 (0.024)	0.004 (0.025)
In couple					0.031 (0.111)	0.031 (0.114)
Respondent is household head					0.118 (0.103)	0.145 (0.110)
Owns home					0.097 (0.070)	0.095 (0.074)
Temptation * 2-day treatment						−0.192** (0.078)
Temptation * 7-day treatment						−0.316** (0.144)
Temptation * 14-day treatment						−0.150 (0.131)
7-day treatment	−0.096 (0.059)	−0.098 (0.060)	−0.099 (0.060)		−0.097 (0.061)	−0.093 (0.072)
14-day treatment	−0.215*** (0.062)	−0.218*** (0.062)	−0.209*** (0.064)		−0.207*** (0.065)	−0.193** (0.079)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	359	359	359	359	359	359
Bootstrapped Replications	1,000	1,000	1,000	1,000	1,000	1,000
Mean of dependent variable	0.783	0.783	0.783	0.783	0.783	0.783
H0: r and β = 0 (p-value)		0.882	0.874	0.877	0.820	0.783

Notes: Bootstrapped standard errors in parentheses (). *** p < 0.01, ** p < 0.05, * p < 0.1. Minimum detectable effects for r and β are shown between [ ].
Time preference parameters used are those summarized in Table 4 on all 359 respondents, for all available price list framing: for both 1,000 CFA and 10,000 CFA, and for 2, 7, 14, 31 and 186 days.
Both time preference parameters are standardized such that the coefficients represent the effect of a 1 standard deviation change in the variable.
Given that β and r (and R) are generated regressors, we use bootstrapping to estimate their standard errors (see Mooney and Duval, 1993).
Our estimation procedure: A cluster bootstrap sample of observations from the 359 individuals is selected, with replacement, from the full sample of 32,310 observations (90 per respondent). Predicted values of β and r are calculated for each respondent, based on the estimated coefficients from equation (2), with these parameters depending on the vector of characteristics described in Section 4.1 (see Table 3). Observations are collapsed to one observation, for each individual in the bootstrapped sample before the second stage probit model is run. The full sample of 32,310 observations is then restored before selecting the second bootstrap sample. Reported standard errors are based on the distribution of the coefficients obtained from 1,000 random bootstrap samples in each model.

tory threshold of 80%. This is described in detail in Table D1. Here, we show that our sample size calculation was powered to detect statistically significant differences for relatively small effect sizes (between 15 and 25 percentage points) for probabilities of keeping the note in the 2-day treatment between 60 and 70%. Partial compliance to treatment, which may have affected experimental design, was not an issue in our case.

4.5. Correlation between hypothetical measures and actual saving behaviour

The key implication of the findings presented in Table 7 is that the MPL-elicited discount rates (based on hypothetical rewards) do not predict the incentivized behaviour in our banknote experiment. However, we recognise that the decision process underlying the banknote experiment does not mirror that in the hypothetical questions. It is a noisy reflection of time preferences and is subject to confounding factors, notably fungibility (which we discuss in Section 5). In Table 8, we consider whether the estimated time preference parameters are significant predictors of whether a respondent comes from a household that owns at least one savings account or participates in a ROSCA. Both variables are measured at the time of the survey, before the MPL questions were administered and before the banknote experiment was introduced.

Models 1 and 2 of Table 8 are analogous to models 3 and 5 of Table 7, with the dependent variable replaced by an indicator of whether the respondent's household owns a savings account (at a bank or MFI), while models 3 and 4 present the same results for ROSCA participation. As there is no underlying reason why random treatment assignment in the banknote experiment should influence either of these outcomes, treatment dummies are excluded in Table 8, although results are fully robust to their inclusion.

As with the banknote results in Table 7, it is not possible to reject the null of no effect of β and r, either jointly or individually. This result holds for both the savings account models and the ROSCA participation models, with and without additional controls. Therefore, the MPL-elicited time preference parameters not only fail to predict incentivized behaviour within the context of the banknote experiment, but also appear uncorrelated with observable economic behaviour relating to saving, which is typically expected to respond to time preferences.

4.6. Robustness checks and statistical power

Within the MPL time preferences elicitation, the share of respondents choosing the sooner hypothetical reward for the two longest delays (1 and 6 months) and for the 1,000 and 10,000

Table 8
Estimated effects of time preference parameters on actual behaviour (probit regressions, marginal effects reported).

	(1)	(2)	(3)	(4)
Dependent Variable:	Savings	Savings	ROSCA	ROSCA
Discount rate (r)	0.089	−0.031	0.003	0.010
	(0.211)	(0.266)	(0.144)	(0.219)
	[0.59]	[0.74]	[0.40]	[0.61]
Present bias parameter (β)	−0.076	−0.050	−0.031	−0.029
	(0.212)	(0.268)	(0.147)	(0.219)
	[0.59]	[0.75]	[0.41]	[0.61]
Risk aversion (R)	0.119	0.173	0.170	0.181
	(0.196)	(0.206)	(0.199)	(0.205)
Income (in 100′000CFA)		0.093**		0.011
		(0.041)		(0.034)
In couple		0.230		−0.029
		(0.144)		(0.130)
Respondent is household head		0.221		−0.019
		(0.148)		(0.127)
Owns home		0.039		0.010
		(0.094)		(0.089)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes
Observations	359	359	359	359
Bootstrapped Replications	1,000	1,000	1,000	1,000
Mean of dependent variable	0.591	0.591	0.393	0.393
H0: r and β = 0 (p-value)	0.913	0.665	0.930	0.978

Notes: Bootstrapped standard errors in parentheses (). *** p < 0.01, ** p < 0.05, * p < 0.1. Minimum detectable effects for r and β are shown between [ ].
The dependent variable Savings takes the value 1 if the respondent comes from a household that owns at least one savings account, in either a bank or microfinance institution (0 otherwise).
The dependent variable ROSCA takes the value 1 if the respondent comes from a household that participates in a ROSCA (0 otherwise).
The inclusion of the treatment assignment variables in the above models does not yield any new interpretation of the results.

CFA amounts are large (for 1,000 CFA: 84 and 96%, respectively, and for 10,000 CFA: 56 and 86%, respectively). To provide a closer comparison to the framing of the banknote experiment, where the longest delay was 14 days, we compute r and β using various subsamples of MPL responses. First, we only use MPL responses for the initial payments of 1,000 CFA. Second, we only use MPL responses for the time delays of 2, 7 or 14 days. Third, we combine both restrictions. Results are presented in columns 1 to 6 of Table E1 and discussed in Appendix E. Estimations are qualitatively similar to those in Table 7.

As a second robustness check, we check the extent to which the results depend on the operationalization of the time preferences. We construct two simple binary variables taking the value of one if an individual gave responses in the MPL consistent with *i)* a lower than median discount rate (and zero otherwise) and *ii)* if an individual gave responses consistent with a higher than median present bias parameter (and zero otherwise). The first variable should be positively correlated with our outcomes of interest, while the second should be negatively correlated. Results are shown in columns 1 and 2 of Table E2. In these estimations, it was, again, impossible to reject the null hypothesis of no effect of the estimated time preference parameters, individually or jointly (at any common significance level), mirroring the main results of our analysis in Table 7. Similarly, we repeat the exercise on the other two outcomes in columns 3–6 of Table E2. Again, we draw the same conclusion. Our estimated time preference parameters do not predict ownership of a savings account or participation in a ROSCA.

Third, we check the extent to which results are affected by the set of controls included in the specification. While the discount rate parameter r decreases in size when controls are included, the present-bias parameter β remains mostly unaffected. Notably, all coefficients remain largely insignificant, regardless of the specification (see Tables 7, 8, E1 and E2).

Fourth, we discuss the extent to which our results suffer from limited statistical power. In doing so, we use the *ex-post* observed standard errors to construct Minimum Detectable Effects (MDEs).[19] We report MDEs in brackets in Table 7, 8 and E2 for the two time preference parameters. These are calculated as 2.8 times the estimated standard error of the relevant coefficients and indicate the minimum effect size our design can detect with 80% power and a 5% significance level.

For comparison, Table A1 summarizes the main features of the papers we included in our review of the use of time preferences in the context of developing countries. We report for each paper the elicitation technique (hypothetical vs incentivized), the operationalization of time preference measures (binary, continuous, categorical) and the significance of the relationship with the outcome.[20] We were unable to find a directly comparable study that employs a continuous (standardized) set of time preference measures to predict a binary outcome (as in our main results in Tables 7 and 8). Hence, to increase comparability, we refer to the MDEs from the simplified version of our analysis in Table E2, where both the outcome and the time preferences are expressed as binary variables.

For papers that use binary variables for both the dependent and independent variables (N = 5), as in Table E2, the average MDE is 24 percentage points. If we only focus on those using hypothetical elicitation methods (N = 4), the average MDE is lower, at 22.5 percentage points. When we consider MDEs as a percentage change with respect to the mean of the dependent variable, for the same pool of studies, the average MDE is 91% (N = 5) and 96.5% (N = 4), respectively. The comparable MDEs from our analysis (the first figure in the [ ] brackets in Table E2) vary between 29 and 33 percentage points for keeping the banknote, between 37

[19] This is discussed in the posts by David McKenzie and Owen Ozier at the World Bank and by Daniel Lakens. The two links are available at: https://blogs.worldbank.org/impactevaluations/why-ex-post-power-using-estimated-effect-sizes-bad-ex-post-mde-not and https://daniellakens.blogspot.com/2014/12/observed-power-and-what-to-do-if-your.html.

[20] More details on Table A1 can be found in Appendix A.

and 41 percentage points for having a savings account and between 32 and 38 percentage points for being a member of a ROSCA. In relative changes with respect to the mean of the dependent variable, these MDEs range between 37 and 42% for keeping the banknote, 63–69% for owning a saving account, 81–97% for ROSCA participation (calculated from Table E2).

In terms of percentage point changes, it is clear that our MDEs are above the average of papers using similar binary measures to capture differences in time discounting. Considering MDEs as a percentage change with respect to the mean of the dependent variable, we are within a similar range. In summary, while not the highest, our MDEs are large, relative to the mean across comparable studies. Importantly, however, none of the papers we use for comparison employ standard error bootstrapping in their estimations, as the time preference indicator variables used in these studies are not derived from a first stage regression. Far simpler measures are used in all cases, commonly an indication of reversal (or consistency) of preference over different time horizons (Ashraf, Karlan, & Yin, 2006; Dupas & Robinson, 2013a; Tarozzi et al., 2014) or grouping respondents based on a specific set of responses. For example, always preferring an earlier reward (Carpenter & Williams, 2014).

To provide a more comparable set of MDEs to these studies, we generate a second set of standard errors and MDEs (shown as the second figure in both sets of brackets in Table E2). Here we continue to split our sample around the median of our time preference parameters, but now treat our first-stage predictions of r and β as observed (rather than estimated). This is done keeping in mind our intention of using the simplest possible method of splitting our respondents into two broad groups, representing those who are relatively more or less impatient.

Based on this second set of standard errors our estimated MDEs vary between 17 and 18 percentage points for keeping the banknote (as the dependent variable), between 18 and 24 points for having a savings account and between 21 and 24 points for being a member of a ROSCA. The MDEs calculated from this second set of non-bootstrapped (robust) standard errors fall well within the range of comparable studies and are aligned to comparable papers finding significant results. With these additional results, we continue to find no correlation between our binary measures of time preferences and our outcomes. The only exception being a minimal effect (at 10%) suggesting a positive correlation between an individual's discount rate and ROSCA participation (Table E2, column 5), which is found not to be robust to the inclusion of the standard control variables in column 6. Thus, even when intentionally presenting our results in this manner – in favour of a rejection of the null - we find, on the whole, no evidence of a significant correlation between either of our elicited time preference measures and our outcomes.

To further establish that the lack of a correlation between our estimated time preference parameters and the three outcomes considered is not due to inflated standard errors (as a result of bootstrapping), we also provide alternative versions of our main Tables 7 and 8 (which employ a continuous measure of both r and β), where we report conventional robust standard errors in place of the original bootstrapped errors (essentially ignoring the first stage variation in the prediction of r and β). These results can be found in Table E3 (replicating Table 7) and Table E4 (replicating Table 8). While we maintain that treating first stage predictions of r and β as observed (constant) rather than predicted variables is not appropriate, when intentionally biasing our results by doing so - towards a rejection of the null - we still find no evidence of a significant correlation between either of our elicited time preference measures and any of our proposed outcomes.

5. Discussion

5.1. Respondent

For 52% of our households, the head of the household was not the respondent. It is plausible that the respondent consulted with their spouse or somebody else in order to decide what to do with the banknote. However, it seemed to us that most respondents dealt with the experiment largely privately and our anecdotal evidence indicates that both spouses manage their income/money independently. A large-scale DHS survey conducted in Senegal (DHS, 2016) also confirms this. Therefore, we would expect the variables indicating gender, and whether the respondent is the head of the household to have no significant effect on the outcome of the experiment, and our results in models 5 and 6 of Table 7 support this. These results also seem to indicate that potential differences in preferences towards keeping the note between those respondents who were married or cohabiting (variable 'In couple') do not play a significant role in our context.

5.2. Fungibility

It is possible to speculate that a household that has better access to liquidity through their own savings or income would have found it easier to keep the 1,000 CFA note. If this were the case, we would expect systematic differences in behaviour among the 'cash-constrained' and the relatively 'cash-abundant' participants within the sample.

Evidence related to the liquidity hypothesis is shown in Table E5 of the appendix. Our results show that there is no significant heterogeneity in the effect of the time preference parameters on keeping the banknote by income quintiles (column 1), ownership of a saving account and (separately) participation in a ROSCA (column 2), or for having access to either saving device (saving account or ROSCA) (column 3). Were any potential effects related to fungibility present, households with higher income should demonstrate greater ease in keeping the note, as should households with access to a saving device. We do not detect any such effect and our results do not seem to support this intuition.

These results suggest that the extent to which liquidity exerted influence on behaviour was negligible. It is likely, however, that this type of behaviour would have been more prevalent, had the experiment been conducted with much larger sums of money. The amount of 1,000 CFA is equivalent to <0.5% of average monthly household income (or 0.61% of its median). However, for the poorest 5% of our sample (who live on 37,500 CFA or less per month), 1,000 CFA represents 2.7% of monthly household income. For the poorest 10% (who live with 62,500 CFA or less), it represents 1.6%. For these households, we argue that 1,000 CFA is not an insignificant amount.[21] Putting the value of the incentive in context, in the 7 papers in Table A1 which use incentivized elicitation strategies, the average reward offered for the short-term horizon was 1.3 USD, which would be equivalent to around 650 CFA in Thies at the time of our survey. This figure excludes papers that are not directly comparable to ours, where high stakes were used but only a few lottery winners receive payment (Tanaka et al. 2010; Becchetti et al. 2017).

Following the discussion in section 2.4, the additional *Questions 4* and *5*, ('Do you plan, yes or no, to do something in order to make sure that you will not spend the note?' and 'If yes what?'), were included (in part) to evaluate any potential bias in behaviour,

[21] To give some additional context, in Thies at the time of the survey 1000 CFA could buy the following items: two or three taxi rides within the town of Thies (depending on the distance covered); a meal in a road-side restaurant for four individuals; two or three pieces of second-hand clothing (t-shirts, cap, sweatpants, etc.).

due to the inherent fungibility of the reward. We wanted to uncover the likelihood that participants in the experiment would consider drawing money from an existing pool of cash or credit, in order to increase their expenditure now, while still managing to retain the specific banknote provided. Our descriptive statistics show that only 0.8% (3 out of 359) of our respondents considered such tactics in order to help them to keep the note, implying this effect is likely to be marginal. In reference to specific potential sources of replacement funds, no individual mentioned the use of savings in ROSCAs, MFIs or bank accounts, when answering *Question 4*.[22] This is not surprising in the case of ROSCAs, given that these informal devices are notoriously inflexible, and are commonly used precisely for the purpose of rendering savings illiquid (Dagnelie & LeMay-Boucher, 2012). Moreover, we argue that it is unlikely that an agent would visit either her bank or MFI office in order to withdraw such a relatively small amount specifically for this purpose (the financial fixed costs of such a transaction alone, would likely represent a significant share of the 1,000 CFA note received).

As we discuss in the following section, some respondents may have answered questions on the banknote experiment in a way that they believe to be socially desirable, and one could also argue that some individuals may not truthfully report their reliance on the fungibility of money to keep the note. Some of our questions are potentially more prone to this issue than others. In particular, Question 3: *'Do you think, yes or no, that you will have difficulties coping with the temptation to spend the banknote?'* may be associated with a certain degree of stigma following a positive response. In the light of the distribution of answers from Question 3, (4 and 5), however (see Section 2.4), where 26% were willing to state that they would have difficulty resisting temptation, we think it's unlikely that respondents would be unwilling to report intentions to draw from an existing pool of cash to avoid spending the specific banknote.

It is also possible that some individuals may have considered the 1,000 CFA banknote as non-fungible even if during our experiment our protocol specifically emphasised that enumerators explain that the money was completely liquid. Our treatment may have nudged some response of mental accounting among individuals, making the possibility of exploiting fungibility less likely. This violation of fungibility by labelling uses to different funds is referred to in Thaler (1990), while Dupas and Robinson (2013a) also make the point that this mental accounting effect can take place without a physical storage place. It could be sufficient for an individual to keep track mentally of what the 1,000 CFA banknote is for, without physically putting it aside.

5.3. Enumerator effects

As noted above, one undesirable effect, which we identified in a small number of our questionnaires during our pilot study, was that some interviewees interpreted being entrusted with the banknote as a test of their trustworthiness in the eyes of the enumerator. Answers suggesting such reputation effects included: 'Because I want to show you (the enumerator) my value', with alternative versions such as: 'I want to show you how I am capable of saving' or 'to show my patience'. We made every effort to eliminate this perception by emphasising that this note was theirs, and that the use they made of it would not be judged or commented upon. Answers to *Question 2*,'Why do you think you can or cannot (keep the money until the specified date)?', suggest that we were able to minimize this effect, as only 1.4% (5 out of 359) of our recipients mentioned anything related to this 'reputation effect' as a potential influence on the decision of whether or not to keep the note.

Although every effort was taken to minimise potential enumerator effects, possible differences in either the methods or style used by the enumerators during the interview (or the characteristics of the enumerators themselves) may have marginally impacted on the outcome of our banknote experiment. We test whether the characteristics of the enumerator influenced their decision to keep the banknote by re-estimating model 6 in Table 7 with the inclusion of indicator variables for each enumerator. The results, shown in column 4 of Table E5, are qualitatively similar to those presented earlier and, following a test of the joint significance of the enumerator variables, while one coefficient (enumerator 2) varied significantly from the base category (enumerator 1), we were unable to reject the null hypothesis of no enumerator effect in a joint test of the coefficients (p-value = 0.641).

6. Conclusion

In this paper, we assess the ability of estimated time preference parameters to predict actual and incentivized saving behaviour. Our measures of time preference are obtained via a commonly applied multiple-price list task, using hypothetical rewards. Intuitively, one would expect to observe a correlation between such elicited measures and decisions based on real money. In our incentivised task, individuals are given a banknote of 1,000 CFA and then informed that if they chose to retain this specific note for a randomly assigned period of time (2, 7 or 14 days) they will receive a second banknote, in effect doubling their initial endowment. It is important to emphasise that our hypothetical and incentivised tasks are not capturing identical decision processes, and that our aim was not to directly compare them. Nevertheless, in our context, one should expect elicited discount parameters to have some explanatory power in predicting the outcomes of our banknote experiment. Even after we consider potential confounding factors (enumerator effects, trust, fungibility and self-control) and control for aversion to risk, this appears not to be the case. Our investigation seems to suggest that hypothetical discounting measures are overall poor predictors of our incentivised behaviour. Furthermore, these hypothetical measures largely uncorrelated with observable economic behaviour related to time preference, such as ownership of a savings account or participation in a ROSCA.

Due to the relative ease of implementation and understanding, the use of hypothetical time preferences elicitation methods is common in survey data collected in developing countries. However, if simple outcomes related to an individual's willingness to delay spending (such as those considered here) provide a reasonable indication of the extent to which an individual discounts the future, the findings of this paper suggest that these elicited variables may in fact be poor predictors of the types of behaviour they are generally believed to capture.

CRediT authorship contribution statement

Jacopo Bonan: Data curation, Funding acquisition, Investigation, Writing – review & editing. **Philippe LeMay-Boucher:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Douglas Scott:** Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[22] A similarly low proportion of answers to *Question 4* (or any related mention of issues linked to fungibility) was also reflected in our pilot survey and led us not to include any retrospective question during our second visit. Thus, our questions related to the banknote experiment are all prospective.

Acknowledgements

We thank Jasmine Wong, Francesca Tamma and Mustapha Diop for their assistance during fieldwork. Thanks are also due to Michel Tenikue, Kyle McNabb, Diego Ubfal and Marc Schaffer. Any remaining errors are our own. LeMay-Boucher wishes to thank the School of Economics at the University of Queensland and the Kellogg Institute for International Studies at the University of Notre Dame for their visiting fellowship programmes. Bonan acknowledges financial support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ ERC grant 342 agreement no. 336155 - project COBHAM "The role of consumer behaviour and heterogeneity in the integrated assessment of energy and climate policies".

Appendix. A. Literature review.

Table A1 reports for each paper (where available), the following information:

Column 2. *Elicitation*: The elicitation technique (hypothetical vs incentivized).

Column 3. *Time preference variables*: The operationalization of the time preferences measures (binary, continuous, categorical).

Column 4. *Correlation with outcomes*: The significance of the relationship with the outcome, and the way the outcome variable is expressed (binary, continuous, categorical). Statistically significant correlations (at a minimum of $p < 0.1$) are signalled when at least one time preference parameter is significantly correlated with at least one outcome considered in the analysis. If the measure of time preference is only reported as the dependent variable in the paper, we report N/A in this column.

Column 5. *MDE ranges (percentage points)*: The MDEs (range and mean) where both the outcome and the time preference measure (s) are binary. The MDEs are expressed in percentage points. Where the time preference measures are the outcome, we report N/A in this column. Where either the dependent variable or the time preference measure is not binary, we report that the MDEs are not directly comparable. For example, this is the case when considering Ashraf, Karlan and Yin (2010), where the dependent variable is categorical. For Column 5, we select the range of MDEs by taking the highest and lowest estimated standard errors for the time preference coefficients, across comparable outcomes (where both the dependent variable and the time preference measures are binary). We also report the mean of the MDEs for each study.

Column 6. *MDE ranges (% change wrt mean of the dependent variable)*: Here, the MDE ranges are also expressed as a percent change with respect to the reported mean of the dependent variable. Again, where the time preference measure is the outcome, we report N/A. Where either the dependent variable or the time preference measure is not binary, we report that the MDE is not directly comparable.

Column 7. *Table and page references*: In each comparable paper, the exact table we use to calculate the MDEs, along with the page reference and the names of the (binary) independent variables for which we calculate these measures are also recorded.

Column 8. *Notes*: These notes include further details, such as values of the mean of the dependent variable or if the time preference measure is the dependent variable.

Our findings are as follows:

1) Our review on the use of time preferences in developing country contexts reports 27 papers. 17 papers use hypothetical elicitations techniques, 8 incentivized measures, one paper uses both, and in one study the elicitation strategy is unclear. Among the 17 papers using hypothetical techniques, 4 consider time preferences only as dependent variables, while 13 studies consider these measures as independent variables. Among incentivized studies, 5 out of 8 use time preferences only as dependent variables. For the papers where the elicitation method was either unclear, or where both real and hypothetical rewards were used, the employed measure of time preference is used as the dependent variable only.

2) Among the 18 studies (13 hypothetical, 5 incentivized) using time preferences as independent variables, 8 found at least one significant correlation with a real-world outcome: 3 from incentivized and 5 from hypothetical elicitation studies.

3) For papers that use binary variables for both the dependent and independent variables (N = 5), the average MDE mean is 24 percentage points. If we only focus on those using hypothetical elicitation methods (N = 4), the average MDE is 22.5 percentage points. When we consider MDE as a percentage change with respect to the mean of the dependent variable, for the same pool of studies, the average MDE is 91% (N = 5) and 96.5% (N = 4), respectively.

4) For papers using hypothetically-elicited time preference measures to predict a real-world outcome (and where results are reported), the majority (7 out 12) do not find a significant correlation between time preferences and the outcome of interest.

Table A1
Papers used in the review of literature.

	Elicitation	Time preference variables	Correlation with outcomes [a]	MDE ranges (percentage points)	MDE ranges (% change wrt mean of the dependent variable) [b]	Table and page references [c]	Notes
Ashraf, Karlan, and Yin (2006)	Hypothetical	Binary	Significant correlation with take-up of a saving product for women (binary)	14–31 (mean = 22)	50–147% (mean = 83%)	Table 5. Page 654 (Variables: Time inconsistent; Impatient now vs 1 month; Patient now vs 1 month; Impatient 6 vs 7 months; Patient 6 vs 7 months; Female*Time inconsistent)	Mean DV: 0.28 (All); 0.31 (Female); 0.21 (Male)
Ashraf, Karlan, and Yin (2010)	Hypothetical	Binary	Significant correlation with women's agency and perceptions of saving behaviour (categorical)	Not directly comparable	Not directly comparable	N/A	
Bauer and Chytilová (2010).	Hypothetical	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable

Table A1 (*continued*)

	Elicitation	Time preference variables	Correlation with outcomes [a]	MDE ranges (percentage points)	MDE ranges (% change wrt mean of the dependent variable) [b]	Table and page references [c]	Notes
Tanaka, Camerer and Nguyen (2010)	Incentivized	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Andersson, Mekonnen, and Stage (2011)	Hypothetical	Continuous	Non-significant correlation with changes in livestock and tree holdings (both continuous)	Not directly comparable	Not directly comparable	N/A	
Bauer, Chytilová, and Morduch (2012)	Incentivized	Binary	Significant correlation among women with self-help group participation (DV1) and borrowing (DV2). Non-significant correlation with other borrowing source (DV3) (all binary)	16–43 (mean = 30)	31–102% (mean = 69%)	Table 5. Page 1131 (Variables: Strongly present-biased; Weakly present-biased; Patient now, impatient in the future)	Mean DV1: 0.652; DV2: 0.426; DV3: 0.281
D'Exelle, Van Campenhout and Lecoutere (2012)	Incentivized	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Shoji et al. (2012)	Hypothetical	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Voors et al. (2012)	Incentivized	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Dupas and Robinson (2013a)	Hypothetical	Binary	Non-significant correlation with calling program officer to open locked savings box in 6 months (DV2) or 12 months (DV3), and ROSCA participation (DV3) (all binary)	10–34 (mean = 20)	25–135% (mean = 76%)	Appendix Table A6. Page 9, (Variables: Present bias; Patient now, impatient later) and Table A8. Page 11 (Variables: Somewhat patient; Present bias; Patient now, impatient later) (online appendix)	Mean DV1: 0.18; DV2: 0.31; DV3: 0.41
Dupas and Robinson (2013b)	Incentivized	Binary	Non-significant correlation with amount saved (continuous)	Not directly comparable	Not directly comparable	N/A	
Blattman, Fiala and Martinez (2014)	Hypothetical (self-reported)	Continuous	Significant correlation with (male) cash earnings not with business assets (both continuous)	Not directly comparable	Not directly comparable	N/A	
Bonan et al. (2019)	Hypothetical	Binary	Significant correlation with ROSCA participation (DV1), not with funeral groups participation (DV2) and significant correlation with either device (DV3) (all binary)	7–21 (mean = 13)	48–135% (mean = 85%)	Table 4 (Panel A). Page 12 (variables: Hyperbolic, Always patient, Female*Hyperbolic)	Mean DV1: 0.15; DV2: 0.14; DV3: N/A
Carpenter and Williams (2014)	Hypothetical	Binary	Non-significant correlation with loan repayment problems (binary)	27–49 (mean = 36)	106–196% (mean = 142%)	Table 6. Page 127 (Variables: patient; impatient)	Mean DV: 0.25
Tarozzi et al. (2014)	Hypothetical	Binary	Non-significant correlation with purchase of anti-malaria bednets (binary)	N/A	N/A	N/A	Coefficients and standard errors are not reported for time preference variables
Rieger (2015)	Incentivized	Categorical	Significant correlation with weight and BMI (both continuous)	Not directly comparable	Not directly comparable	N/A	
Carvalho et al. (2016)	Hypothetical	Binary	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Ubfal (2016)	Hypothetical and incentivized	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Becchetti, Castriota and Conzo (2017)	Incentivized	Continuous	Significant correlation with share of endowment sent or expected in a dictator	Not directly comparable	Not directly comparable	N/A	

(*continued on next page*)

Table A1 (*continued*)

	Elicitation	Time preference variables	Correlation with outcomes [a]	MDE ranges (percentage points)	MDE ranges (% change wrt mean of the dependent variable) [b]	Table and page references [c]	Notes
			game (continuous)				
Cassar, Healy and Von Klesser (2017)	Incentivized	Continuous	Significant correlation with amount of money sent in a trust game	Not directly comparable	Not directly comparable	N/A	
Hoffmann and Muttarak (2017)	Hypothetical	Categorical	Non-significant correlation with disaster preparedness (binary)	Not directly comparable	Not directly comparable	N/A	
Le Cotty et al. (2018)	Hypothetical	Continuous	Significant correlation with fertiliser use and maize area (both continuous)	Not directly comparable	Not directly comparable	N/A	
Maertens and Vande Velde (2017)	Hypothetical	Binary	Non-significant correlation with income and agricultural production outcomes (continuous)	Not directly comparable	Not directly comparable	N/A	
Nguyen and Rieger (2017)	Unclear	Binary	N/A	N/A	N/A	N/A	Time preference is the dependent variable
Perez-Arce (2017)	Hypothetical	Categorical	Non-significant correlation between patience for money and smoking (binary) and hours worked per week (continuous)	Not directly comparable	Not directly comparable	N/A	
Alemayehu, Beuving and Ruben (2018)	Hypothetical	Continuous	N/A	N/A	N/A	N/A	Results are not shown
Berry, Karlanand Pradhan (2018)	Hypothetical	Continuous	N/A	N/A	N/A	N/A	Time preference is the dependent variable

Notes: [1] A significant correlation with outcomes refers to a correlation of at least p < 0.1.
[2]MDE ranges with respect to the mean of the dependent variable calculated as: MDE in percentage points (p.p.) / mean of the dependent variable.
[3]The listed table (or tables) relates to the table used to calculate the MDEs, with the corresponding page numbers. The variables listed were those used to calculate the MDEs.

Appendix B. . Preferences elicitation.

Table B1.

Table B2.
Table B3.

Table B1
Eliciting the time preference parameters.

Panel A: Amount proposed for today 10,000 CFA		
A	B	A or B?
Today	In 2 days	
10,000	10,000	
10,000	10,500	
10,000	11,000	
10,000	12,500	
10,000	15,000	
10,000	17,500	
10,000	20,000	
10,000	25,000	
10,000	30,000	
Panel B: Amount proposed for today 1,000 CFA		
A	B	A or B?
Today	In 2 days	
1,000	1,000	
1,000	1,050	
1,000	1,100	
1,000	1,250	
1,000	1,500	
1,000	1,750	
1,000	2,000	
1,000	2,500	
1,000	3,000	

Notes: Four additional sets of choices were offered where the values in A and B were identical, but the time delay was 7, 14 days, 1 months and 6 months.

Table B2
Violations of time preference transitivity in the pilot study.

	(1)	(2)	(3)
	N	Mean	SD
Panel A. Inconsistencies in time preference elicitation			
2 days - high stakes	60	0	0
2 weeks - high stakes	60	0	0
1 month - high stakes	60	0.017	0.129
6 months - high stakes	60	0	0
2 days - low stakes	60	0	0
2 weeks - low stakes			
1 month - low stakes	60	0	0
6 months - low stakes	60	0	0
Any inconsistency	60	0.017	0.129
Panel B. Inconsistencies in risk preference elicitation			
low stakes - gains game	60	0.033	0.181
Inconsistency in high stakes - gains game	60	0.033	0.181
Inconsistency in low stakes - loss game	60	0	0
Inconsistency in high stakes - loss game	60	0.017	0.129
Any inconsistency	60	0.05	0.219

Table B3
Eliciting risk preferences.

Number Marbles type1	Number Marbles type2	Lottery A		Lottery B	
		Successful payoff	Unsuccessful payoff	Successful payoff	Unsuccessful payoff
8	2	600	0	200	100
7	3	600	0	200	100
6	4	600	0	200	100
5	5	600	0	200	100
4	6	600	0	200	100
3	7	600	0	200	100

Appendix C. . Alternative discounting functions.

In Table C1, we report the estimated time preference parameters from the four nested discounting functions contained within equation (1). This equation follows Benhabib, Bisin & Schotter (2010). Column 1 of the table indicates our estimate of the discount rate under the assumption of exponential discounting (e^{-rt}), whereby the parameters in equation (1) are restricted, such that $\beta = 1$ and θ is approaching 1. Column 2 shows the estimated discount rate under pure hyperbolic discounting ($1/(1 + rt)$), with $\beta = 1$ and $\theta = 2$, in equation (1). Column 3 represents the time preference parameter estimates under the assumption of quasi-hyperbolic discounting, where the discount factor takes the form βe^{-rt} and individuals discount all non-immediate payments by a factor β. Finally, column 4 reports the results of the full model, imposing no restriction on the parameters in equation (1).

The data contains 32,310 (pooled) observations (90 questions, for each of the 359 individuals). The estimated parameters of the discount factor are broadly comparable to those in Tanaka et al. (2010). The adjusted R^2 for the models suggests that the full model, with non-restricted values for Theta (θ), seems to perform only marginally better than the quasi-hyperbolic specification. Furthermore, a test of the hypothesis that $\theta = 2$ in the full model (the assumption required for pure hyperbolic discounting) is strongly rejected (p-value = 0.00). We also reject the null that $\beta = 1$ in the last 2 models (with p-values close to 0). These results seem to indicate that the quasi-hyperbolic model is a strong candidate. It also allows comparison with previous papers (e.g., Tanaka et al. (2010)).

Appendix D. . Ex-Ante power of the banknote experiment.

We work with a sample of 120 for both the treatment group (N2) and the control group (N1). N1 is the group of individuals treated

Table C1
Estimated time preference parameters: four alternative discount functions (all observations pooled).

	(1) Exponential	(2) Hyperbolic	(3) Q-hyperbolic	(4) Full model
Discount rate (r)	0.078***	0.139***	0.043***	0.309***
	(0.000)	(0.000)	(0.000)	(0.014)
Beta (β)			0.733***	0.945***
			(0.002)	(0.008)
Theta (θ)				4.281***
				(0.025)
Noise parameter (μ)	−3.424***	−5.278***	−5.178***	−5.790***
	(0.085)	(0.146)	(0.135)	(0.149)
Observations	32,310	32,310	32,310	32,310
Adjusted R^2	0.401	0.450	0.464	0.465

Notes: Individual-level, cluster-robust standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

Table D1
Ex-anti power computation (7-day treatment).

Panel A: (subsamples: N1 = N2 = 120)				
Prob. of keeping note P1(in % points)	Prob. of keeping note P2(in % points)	Effect size (in % points)	Power (%) *α set at 5%*	Power (%) *α set at 10%*
70	50	20	94	97
70	45	25	99	99
70	40	30	99	99
65	50	15	76	86
65	45	20	93	97
65	40	25	99	99
60	45	15	75	85
60	40	20	93	97

with the 2-day offer. We show the values of power for three initial values of the probability of keeping the note for the control group (P1): 70, 65 and 60% rates. These figures are taken from our pilot (before our survey) and appeared to us to be the most plausible. During our design phase, we expected an effect size of around 15–35 percentage points for the 7-day treatment and around 30–50 percentage points for the 14-day treatment. The discussion on effect size related to Table D1 refers to percentage points (p.p.).

For an expected probability of retention of the note of 70% for the control group (P1): we could detect an effect size of 15p.p. (13p.p.) or larger with a power above 80% for a significance level α set at 5% (10%). So, for an anticipated effect size of 20 or above, we have at least a power of 94%. For an expected probability of retention of the note of 65% for the control group (P1): we could detect an effect size of 16p.p. (13.5p.p.) or larger with a power above 80% for a significance level α set at 5% (10%). So, for an anticipated effect size of 20 or above, we have at least a power of 93%. The figures are similar if we use a probability of retention of the note of 60p.p. for the control group (P1).

Overall, our RCT components were not weakened by the size of our samples. A power of 80% or above is often considered as a minimum threshold and we do get to that level or above for most effect sizes that are below the expected size we anticipated in our design phase. We do have power values slightly below that level (at 76% and 75%) only for small size effects of 15 percentage points and a significance level α set at 5%. For all other cases, the power is significantly above this 80% threshold. The computation of the power of our test on the coefficient of 7-day treatment is not related to the second orthogonal treatment 14-day. As we can deduct from Table D1, an even smaller sample size than we have (120) is required to have enough power for the 14-day treatment, which has a larger expected effect size.

Appendix E. . Robustness checks and further results.

Relative to the treatment durations in the banknote experiment, the MPL questions contained a larger range of delays between the initial and the later reward (2, 7, 14, 31 and 186 days) and also considered two initial amounts (1,000 and 10,000 CFA).

Table E1
Robustness tests of the relationship between keeping the banknote in our experiment and the estimated time preference parameters derived from the MPL questions.

	(1)	(2)	(3)	(4)	(5)	(6)
	The dependent variable is Banknote kept = 1 (=0 otherwise)					
MPL questions:	1,000 CFA		Delay < 14 days		1,000 CFA and < 14 days	
Discount rate (r)	−0.039	−0.099	0.070	0.007	−0.044	−0.099
	(0.064)	(0.100)	(0.077)	(0.114)	(0.069)	(0.101)
Present bias (β)	0.049	0.118	−0.064	−0.055	0.052	0.119
	(0.059)	(0.084)	(0.078)	(0.114)	(0.061)	(0.087)
Risk aversion (R)	−0.182	−0.167	−0.184	−0.170	−0.179	−0.165
	(0.157)	(0.158)	(0.160)	(0.161)	(0.157)	(0.158)
Income (in 100′000CFA)		0.026		0.004		0.024
		(0.034)		(0.024)		(0.035)
In couple		0.036		0.030		0.037
		(0.099)		(0.120)		(0.100)
Respondent is head		0.118		0.143		0.118
		(0.081)		(0.119)		(0.087)
Owns home		0.065		0.095		0.065
		(0.072)		(0.074)		(0.072)
Temptation*2-day treatment		−0.199**		−0.191**		−0.200**
		(0.078)		(0.078)		(0.078)
Temptation*7-day treatment		−0.344**				
		(0.143)				
Temptation*14-day treatment		−0.129				
		(0.129)				
7-day treatment		−0.087	−0.099	−0.092	−0.093	−0.087
		(0.072)	(0.060)	(0.073)	(0.061)	(0.072)
14-day treatment		−0.201**	−0.209***	−0.193**	−0.209***	−0.200**
		(0.079)	(0.064)	(0.079)	(0.064)	(0.079)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	359	359	359	359	359	359
Bootstrapped Replications	1,000	1,000	1,000	1,000	1,000	1,000
Mean of dependent variable	0.783	0.783	0.783	0.783	0.783	0.783
H0: r and β = 0 (p-value)	0.969	0.334	0.666	0.661	0.689	0.348

Notes: Bootstrapped standard errors in parentheses (). *** p < 0.01, ** p < 0.05, * p < 0.1. Minimum detectable effects for r and β are shown between [ ].

Table E2
Estimated effects of dichotomous time preference variables on the probability of keeping the banknote and actual behaviour (probit regressions, marginal effects reported).

Dependent Variable:	(1) Banknote kept	(2) Banknote kept	(3) Savings	(4) Savings	(5) ROSCA	(6) ROSCA
Below median r = 1	0.004	0.064	−0.006	0.107	0.143	0.134
	(0.104; 0.061)	(0.117; 0.063)	(0.137; 0.073)	(0.147; 0.081)	(0.116; 0.074*)	(0.132; 0.084)
	[0.29; 0.17]	[0.33; 0.18]	[0.38; 0.18]	[0.41; 0.23]	[0.32; 0.21]	[0.37; 0.24]
Above median β = 1	−0.002	−0.071	0.049	0.045	0.103	0.128
	(0.104; 0.061)	(0.117; 0.062)	(0.131; 0.072)	(0.144; 0.084)	(0.121; 0.072)	(0.137; 0.079)
	[0.29; 0.17]	[0.33; 0.17]	[0.37; 0.20]	[0.40; 0.24]	[0.34; 0.20]	[0.38; 0.22]
Risk aversion (R)	−0.176	−0.170	0.151	0.206	0.220	0.235
	(0.159; 0.134)	(0.159; 0.133)	(0.194; 0.178)	(0.204; 0.183)	(0.198; 0.184)	(0.206; 0.185)
Income (in 100′000 CFA)		0.006		0.097		0.017
		(0.021; 0.016)		(0.030***; 0.022***)		(0.027; 0.021)
In couple		0.030		0.152		−0.043
		(0.098; 0.084)		(0.112; 0.102)		(0.112; 0.104)
Respondent is household head		0.161		0.130		−0.040
		(0.079**; 0.071**)		(0.099; 0.090)		(0.093; 0.091)
Owns home		0.096		0.027		−0.004
		(0.074; 0.060)		(0.093; 0.074)		(0.089; 0.076)
Temptation * 2-day treatment		−0.205				
		(0.078***; 0.062***)				
Temptation * 7-day treatment		−0.302				
		(0.144**; 0.130**)				
Temptation * 14-day treatment		−0.132				
		(0.129; 0.119)				
7-day treatment	−0.096	−0.093				
	(0.061: 0.055*)	(0.073; 0.055*)				
14-day treatment	−0.206	−0.194				
	(0.064***; 0.055***)	(0.079**; 0.055***)				
Neighbourhood fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	359	359	359	359	359	359
Mean of dependent variable	0.783	0.783	0.591	0.591	0.393	0.393
H0: r and β = 0 (p-value)	0.997	0.444	0.899	0.762	0.462	0.538

Notes: Standard errors (a; b) and MDEs [a; b] reported in parentheses. Standard errors (a) and MDEs [a] are derived from estimations obtained from 1,000 bootstrapped replications. Heteroskedasticity robust standard errors (b) and MDEs [b] are derived from estimations treating first-stage predictions of r and β as observed values, and are thus not bootstrapped, to allow comparison with similar studies.
Test of H0: r and β = 0 are based on bootstrapped results. *** p < 0.01, ** p < 0.05, * p < 0.1.
The dependent variable Banknote kept takes the value 1 if the respondent retained the banknote until the second visit (0 otherwise).
The dependent variable Savings takes the value 1 if the respondent comes from a household which owns at least one savings account, in either a bank or microfinance institution (0 otherwise).
The dependent variable ROSCA takes the value 1 if the respondent comes from a household which participates in a ROSCA (0 otherwise).
The inclusion of the treatment assignment variables in the above models does not yield any new interpretation of the results.

In columns (1) and (2) of Table E1 we restrict the sample of MPL questions used to calculate the discount parameters r and β to only those questions relating to the initial amount of 1,000 CFA, in line with the 1,000 CFA banknote used in the experiment. In columns (3) and (4) we restrict the MPL questions used to only those relating to a time between (hypothetical) payments of 14 days or less. Again, this is intended to create a closer comparison to the conditions of the banknote experiment, where respondents were required to keep the note for no more than 14 days. Columns (5) and (6) impose both restrictions on the MPL questions used.

Finally, in columns (7) and (8) dichotomous variables are used in place of the continuous variables r and β. The first of these new variables takes the value 1 if a respondent's estimated discount rate is below the sample median for r. The second variable takes the value 1 if the estimated present bias parameter is above the sample median for β. Both a low discount rate and a high present bias parameter should indicate a higher degree of patience and, therefore, a higher probability of retaining the banknote.

Table E1 indicates that even when matching the initial reward offered or the period required to wait for the future reward, the MPL-derived discount parameters remain poor predictors of the outcome of the banknote experiment. Similarly, employing a dichotomous variable in place of the estimated r and β yields no evidence of a significant effect on the banknote experiment (individually or jointly).

Table E2 replicates Table 7 (columns 3 and 6) and Table 8, using two binary variables to represent the discount rate and present bias coefficient, defined as below the median value of the estimated parameter r and above the median value of the estimated present bias parameter β.

In addition to providing a test of the robustness of our results to a simpler representation of time preference, these results also provide a set of MDEs which are directly comparable to a number of papers in our literature review (Table A1): using binary measures of time preference as well as a binary dependent variable.

We provide two sets of standard errors and MDEs (for r and β). The first set are derived from estimations obtained from 1,000 bootstrapped replications. The second set of standard errors and MDEs are derived from estimations treating first-stage predictions of r and β as observed values, and are thus not bootstrapped, to allow comparison with similar studies which do not employ this treatment of the errors (see Section 4.6).

Odd numbered columns in Table E2 display the results without individual controls, while even numbered columns include these controls. In both cases neighbourhood fixed effects are included.

Table E3 and E4 report versions of Tables 7 and 8, respectively. In the tables below the standard errors are no longer bootstrapped and instead conventional (Huber/White) heteroskedasticity robust errors are reported. In both tables we find no evidence of a significant correlation between either of our time preference parameters and any of our outcomes.

It should be stated clearly that we do not believe this treatment of the errors is appropriate, as in doing so we would be ignoring the

Table E3
Estimated effects of time preference parameters on the probability of keeping the banknote, treating first-stage predictions of r and β as observed values (probit regressions, marginal effects reported).

	(1)	(2)	(3)	(4)	(5)	(6)
	The dependent variable is Banknote kept = 1 (=0 otherwise)					
Discount rate (r)		0.061	0.064	0.058	0.009	0.002
		(0.051)	(0.051)	(0.051)	(0.072)	(0.072)
		[0.14]	[0.14]	[0.14]	[0.20]	[0.20]
Present bias parameter (β)		−0.053	−0.058	−0.052	−0.048	−0.051
		(0.049)	(0.049)	(0.050)	(0.051)	(0.051)
		[0.14]	[0.14]	[0.14]	[0.14]	[0.14]
Risk aversion (R)			−0.184	−0.267*	−0.141	−0.170
			(0.143)	(0.145)	(0.144)	(0.145)
Income (in 100′000CFA)					0.011	0.004
					(0.016)	(0.016)
In couple					0.031	0.031
					(0.098)	(0.098)
Respondent is household head					0.118	0.145
					(0.101)	(0.101)
Owns home					0.097*	0.095*
					(0.057)	(0.057)
Temptation * 2-day treatment						−0.192***
						(0.063)
Temptation * 7-day treatment						−0.316**
						(0.132)
Temptation * 14-day treatment						−0.150
						(0.124)
7-day treatment	−0.096*	−0.098*	−0.099*		−0.097*	−0.099*
	(0.058)	(0.058)	(0.058)		(0.058)	(0.058)
14-day treatment	−0.215***	−0.218***	−0.209***		−0.207***	−0.209***
	(0.055)	(0.055)	(0.055)		(0.056)	(0.055)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	359	359	359	359	359	359
Mean of dependent variable	0.783	0.783	0.783	0.783	0.783	0.783
H0: r and β = 0 (p-value)		0.882	0.874	0.877	0.820	0.783

Notes: Heteroskedasticity robust standard errors in parentheses (). *** p < 0.01, ** p < 0.05, * p < 0.1.
Minimum detectable effects for r and β are shown between [ ].
Both discount parameters are standardized such that the coefficients represent the effect of a 1 standard deviation change in the variable.

Table E4
Estimated effects of time preference parameters on actual behaviour, treating first-stage predictions of r and β as observed values (probit regressions, marginal effects reported).

	(1) Savings	(2) Savings	(3) ROSCA	(4) ROSCA
Discount rate (r)	0.089	−0.031	0.003	0.010
	(0.064)	(0.093)	(0.066)	(0.091)
	[0.18]	[0.26]	[0.18]	[0.25]
Present bias parameter (β)	−0.076	−0.050	−0.031	−0.029
	(0.065)	(0.068)	(0.066)	(0.068)
	[0.18]	[0.19]	[0.18]	[0.19]
Risk aversion (R)	0.119	0.173	0.170	0.181
	(0.173)	(0.181)	(0.171)	(0.172)
Income (in 100′000CFA)		0.093***		0.011
		(0.021)		(0.019)
In couple		0.230*		−0.029
		(0.130)		(0.125)
Respondent is household head		0.221		−0.019
		(0.135)		(0.129)
Owns home		0.039		0.010
		(0.076)		(0.074)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes
Observations	359	359	359	359
Mean of dependent variable	0.591	0.591	0.393	0.393
H0: r and β = 0 (p-value)	0.378	0.355	0.611	0.874

Notes: Heteroskedasticity robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.
Minimum detectable effects for r and β are between [ ].
The dependent variable Savings takes the value 1 if the respondent comes from a household which owns at least one savings account, in either a bank or microfinance institution (0 otherwise).
The dependent variable ROSCA takes the value 1 if the respondent comes from a household which participates in a ROSCA (0 otherwise).
The inclusion of the treatment assignment variables in the above models does not yield any new interpretation of the results.

variance attributable to our first-stage estimates of r and β. However, the results in Table E3 and E4 should provide reassurance that our inability to detect an effect of the time preference parameters, in the original tables 7 and 8, is not purely a result of larger standard errors obtained through bootstrapping.

Table E5
Fungibility and enumerator effects.

	(1)	(2)	(3)	(4)
	The dependent variable is Banknote kept = 1 (=0 otherwise)			
7-day treatment	−0.138***	−0.104*	−0.098	−0.130**
	(0.047)	(0.063)	(0.063)	(0.062)
14-day treatment	−0.246***	−0.217***	−0.210***	−0.238***
	(0.065)	(0.070)	(0.068)	(0.071)
Discount rate (r)	−0.115	0.039	0.038	0.043
	(0.429)	(0.195)	(0.192)	(0.125)
Present bias parameter (β)	0.110	−0.049	−0.049	−0.045
	(0.406)	(0.195)	(0.192)	(0.128)
Risk aversion (R)	0.164	−0.151	−0.161	−0.158
	(0.188)	(0.162)	(0.165)	(0.152)
Income quintile 1 (lowest)	−0.061			
	(0.109)			
Income quintile 2	0.067			
	(0.135)			
Income quintile 4	−0.031			
	(0.083)			
Income quintile 5 (highest)	0.007			
	(0.211)			
Discount rate*Income quintile 1	0.238			
	(0.704)			
Discount rate*Income quintile 2	0.225			
	(0.799)			
Discount rate*Income quintile 4	0.274			
	(0.427)			
Discount rate*Income quintile 5	0.029			
	(0.566)			
Present bias*Income quintile 1	−0.137			
	(0.726)			
Present bias*Income quintile 2	−0.260			
	(0.779)			
Present bias*Income quintile 4	−0.306			
	(0.417)			
Present bias*Income quintile 5	−0.051			
	(0.580)			
Savings account		0.068		
		(0.097)		
Member of ROSCA		−0.012		
		(0.096)		
Discount rate*Savings account		−0.248		
		(0.330)		
Present bias*Savings account		0.166		
		(0.334)		
Discount rate*ROSCA		0.211		
		(0.288)		
Present bias*ROSCA		−0.176		
		(0.294)		
	(1)	(2)	(3)	(4)
	The dependent variable is Banknote kept =1 (=0 otherwise)			
Any savings device			0.052	
			(0.059)	
Discount rate*Any savings device			−0.084	
			(0.215)	
Present bias*Any savings device			0.025	
			(0.220)	
Enumerator 2				0.928**
				(0.453)
Enumerator 3				−0.379
				(0.828)
Enumerator 4				−0.001
				(0.468)
Enumerator 5				−0.159
				(0.597)
Enumerator 6				−0.243
				(0.769)
Enumerator 7				−0.078
				(0.749)
Enumerator 9				0.098
				(0.554)
Enumerator 10				−0.586
				(0.848)
Enumerator 11				−0.444

(*continued on next page*)

Table E5 (*continued*)

	(1)	(2)	(3)	(4)
	The dependent variable is Banknote kept =1 (=0 otherwise)			
				(0.903)
Neighbourhood fixed effects	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Observations	359	359	359	359
Bootstrapped Replications	1,000	1,000	1,000	1,000
Mean of dependent variable	0.783	0.783	0.783	0.783
H0: All Enumerator effects = 0 (p-value)				0.641

Notes: Bootstrapped standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Where not already reported, control variables include Income, In couple, Respondent is household head, and Owns home.

References

Alemayehu, M., Beuving, J., & Ruben, R. (2018). Risk Preferences and Farmers' Livelihood Strategies: A Case Study from Eastern Ethiopia. *Journal of International Development, 30*, 1369–1391.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9*(4), 383–405.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583–618.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting Behavior: A Reconsideration. *European Economic Review, 71*, 15–33.

Andersson, C., Mekonnen, A., & Stage, J. (2011). Impacts of the Productive Safety Net Program in Ethiopia on livestock and tree holdings of rural households. *Journal of Development Economics, 94*, 119–126.

Andreoni, J., & Sprenger, C. (2012). Risk preferences are not time preferences. *The American Economic Review, 102*(7), 3357–3376.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics, 121*(2), 635–672.

Ashraf, N., Karlan, D., & Yin, W. (2010). Female Empowerment: Impact of a Commitment Savings Product in the Philippines. *World Development, 38*(3), 333–344.

Atkeson, A., & Ogaki, M. (1996). Wealth-varying intertemporal elasticities of substitution: Evidence from panel and aggregate data. *Journal of Monetary Economics, 38*(3), 507–534.

Bauer, M., & Chytilová, J. (2010). The impact of education on subjective discount rate in Ugandan villages. *Economic Development and Cultural Change, 58*(4), 643–669.

Bauer, M., Chytilová, J., & Morduch, J. (2012). Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India. *American Economic Review, 102*, 1118–1139.

Becchetti, L., Castriota, S., & Conzo, P. (2017). Disaster, Aid, and Preferences: The Long-run Impact of the Tsunami on Giving in Sri Lanka. *World Development, 94*, 157–173.

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior, 69*(2), 205–223.

Berry, J., Karlan, D., & Pradhan, M. (2018). The Impact of Financial Education for Youth in Ghana. *World Development, 102*, 71–89.

Blattman, C., Fiala, N., & Martinez, S. (2014). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *The Quarterly Journal of Economics, 129*, 697–752.

Bonan, J., McNabb, K., LeMay-Boucher, P., & Tomavo, C. (2019). Time preferences and commitment devices: Evidence from ROSCAs and funeral groups in Benin. *Oxford Development Studies, 40*(3), 356–372.

Carpenter, J., & Williams, T. (2014). Peer Monitoring and Microcredit: Field Experimental Evidence from Paraguay. *Oxford Development Studies, 42*(1), 111–135.

Carvalho, L. S., Prina, S., & Sydnor, J. (2016). The effect of saving on risk attitudes and intertemporal choices. *Journal of Development Economics, 120*, 41–52.

Cassar, A., Healy, A., & von Kessler, C. (2017). Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand. *World Development, 94*, 90–105.

Coller, M., & Williams, M. (1999). Eliciting individual discount rates. *Experimental Economics, 2*, 107–127.

Dagnelie, O., & LeMay-Boucher, P. (2012). Rosca participation in Benin: A commitment issue. *Oxford Bulletin of Economics and Statistics, 74*(2), 235–252.

D'Exelle, B., van Campenhout, B., & Lecoutere, E. (2012). Modernisation and Time Preferences in Tanzania: Evidence from a Large-Scale Elicitation Exercise. *Journal of Development Studies, 48*(4), 564–580.

Dhs (2016). *Sénégal : Enquête Démographique et de Santé Continue (EDS-Continue) 2015*. USA: DHS Program.

Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *American Economic Review, 101*(6), 2350–2390.

Dupas, P., & Robinson, J. (2013a). Why Don't the Poor Save More? Evidence from Health Savings Experiments. *American Economic Review, 103*(4), 1138–1171.

Dupas, P., & Robinson, J. (2013b). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics, 5*(1), 163–192.

Harrison, G., Lau, M., & Williams, M. (2002). Estimating individual discount rates in Denmark: A field experiment. *American Economic Review, 92*(5), 1606–1617.

Harrison, G. W., Lau, M. I., Rutström, E. E., & Sullivan, M. B. (2005). Eliciting risk and time preferences using field experiments: Some methodological issues. *Research in Experimental Economics, 10*, 125–218.

Harrison, G. W., Humphrey, S. J., & Verschoor, A. (2010). Choice under uncertainty: Evidence from Ethiopia. *India and Uganda. The Economic Journal, 120*(543), 80–104.

Hinvest, N. S., & Anderson, I. M. (2010). The effects of real versus hypothetical reward on delay and probability discounting. *The Quarterly Journal of Experimental Psychology, 63*(6), 1072–1084.

Hoffmann, R., & Muttarak, R. (2017). Learn from the Past, Prepare for the Future: Impacts of Education and Experience on Disaster Preparedness in the Philippines and Thailand. *World Development, 96*, 32–51.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effect. *American Economic Review, 92*(5), 1644–1655.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77*(2), 129–146.

Kirby, K. N., & Maraković, N. (1995). Modeling myopic decisions: Evidence for hyperbolic delay-discounting with subjects and amounts. *Organization Behavior and Human Decision Processes., 64*(1), 22–30.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443–477.

Lawrance, E. C. (1991). Poverty and the rate of time preference: Evidence from panel data. *Journal of Political economy, 99*(1), 54–77.

Le Cotty, T., Maître d'Hôtel, E., Soubeyran, R., & Subervie, J. (2018). Linking Risk Aversion, Time Preference and Fertiliser Use in Burkina Faso. *The Journal of Development Studies, 54*, 1991–2006.

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Experimental and Clinical Psychopharmacology, 11*(2), 139–145.

Maertens, M., & Vande Velde, K. (2017). Contract-farming in Staple Food Chains: The case of rice in Benin. *World Development, 95*, 73–87.

Meier, S., & Sprenger, C. (2010). Present-biased preferences and credit card borrowing. *American Economic Journal: Applied Economics, 2*, 193–210.

Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping: A Nonparametric Approach to Statistical Inference*. Newbury Park, CA: Sage.

Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of Experimental Analysis of Behavior, 63*(3), 263–276.

Nguyen, T. C., & Rieger, M. (2017). Community-driven development and social capital: Evidence from Morocco. *World Development, 91*, 28–52.

Perez-Arce, F. (2017). The effect of education on time preferences. *Economics of Education Review, 56*, 52–64.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior, 55*(2), 233–244.

Rieger, M. (2015). Risk aversion, time preference and health production: Theory and empirical evidence from Cambodia. *Economics and Human Biology, 17*, 1–15.

Shoji, M., Aoyagi, K., Kasahara, R., Sawada, Y., & Ueyama, M. (2012). Social Capital Formation and Credit Access: Evidence from Sri Lanka. *World Development, 40*, 2522–2536.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review, 100*(1), 557–571.

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., & Yoong, J. (2014). Micro-loans, insecticide-treated bednets, and malaria: Evidence from a randomized controlled trial in Orissa, India. *American Economic Review, 104*, 1909–1941.

Thaler, R. H. (1990). Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives, 4*(1), 193–205.

Ubfal, D. (2016). How general are time preferences? Eliciting good-specific discount rates. *Journal of Development Economics, 118*, 150–170.

Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., Lensink, R., & Van Soest, D. P. (2012). Violent conflict and behavior: A field experiment in Burundi. *American Economic Review, 102*, 941–964.