

# Magnitude Effect in Intertemporal Allocation Tasks

---

**Chen Sun** (HU Berlin)  
**Jan Potters** (Tilburg University)

Discussion Paper No. 159

June 24, 2019

# Magnitude Effect in Intertemporal Allocation Tasks

Chen Sun and Jan Potters\*

March 7, 2019

## Abstract

We investigate how intertemporal allocation of monetary rewards is influenced by the size of total budget, with a particular interest in the channels of influence. We find a significant magnitude effect: the budget share allocated to the later date increases with the size of the budget. At the aggregate level as well as at the individual level, we find magnitude effects both on the discount rate and on intertemporal substitutability (i.e. utility curvature). The latter effect is consistent with theories in which the degree of asset integration is increasing in the stake.

Keywords: time preference, magnitude effect, Convex Time Budget method

JEL-code: C91, D12, D91

\* Sun: Humboldt-Universität zu Berlin (e-mail: chen.sun@hu-berlin.de); Potters: Department of Economics and CenterER, Tilburg University (e-mail: j.j.m.potters@uvt.nl). We are grateful for the helpful discussions with Arthur van Soest, and the insightful comments of Mohammed Abdellaoui, Elena Cettolin, Stephen Cheung, Mark Dean, Sebastian Ebert, Glenn Harrison, Gijs van de Kuilen, Guy Mayraz, Bin Miao, Wieland Müller, Kirsten Rohde, Karl Schlag, Stefan Trautmann, Yuehui Wang, Yuxin Yao, Songfa Zhong, and participants at the Tilburg mEETing, the Xiamen Workshop on Experimental Economics 2014, the ENTER Jamboree 2015, the Tilburg Economics Workshop, the MBEES 2015, the EEA Annual Congress 2015, the ESA European meeting 2015 and the IMEBESS 2016. We thank the Netherlands Organization of Scientific Research (NWO) for financial support. Chen Sun further acknowledges financial support from the German Science Foundation (Deutsche Forschungsgemeinschaft) through the project CRC TRR 190.

The prediction of the standard consumption-savings model, that people always discount an income at the market interest rate, has been found to be inconsistent with empirical results.<sup>1</sup> One important anomaly, dating back to Thaler (1981), is the magnitude effect: when comparing a smaller sooner reward with a larger later reward, people favor the later reward more often as the amounts of the two rewards are scaled up. Studies on the magnitude effect help us better understand the relation between intertemporal choices in the lab and those in the real life, given the fact that most lab experiments use small stakes but many choices in the real life involve large stakes.

While a few experiments observed the magnitude effect, little effort is made to explore its mechanisms: are people more patient to larger amounts of money, or do they find larger amounts of money to be more substitutable across time? Both mechanisms account for the magnitude effect in a *single-reward task* where a decision-maker can only receive a reward on a single date (sooner or later), but they have different meanings and predict differently in a more general situation. Ignorance about the mechanisms makes us unable to predict intertemporal choices in general situations. In the meantime, there is not yet any empirical work on how intertemporal choices from a set of cash flows (rather than rewards on a single date) change as all amounts are scaled up.

In this paper, we perform a lab experiment to investigate how choices in intertemporal allocation tasks change with the magnitude of rewards, and in particular, whether the magnitudes impact intertemporal preferences through patience (the discount rate) or through intertemporal substitutability (the atemporal utility function).

<sup>1</sup> To be more precise, people discount at the market rate unless they do not have enough liquidity to smooth consumption between relevant periods.

Several experiments on time preferences have reported a magnitude effect.<sup>2</sup> Though most early studies are based on hypothetical decisions, there are also some real-stake experiments that found a magnitude effect (Holcomb and Nelson 1992; Kirby 1997; Kirby, Petry and Bickel 1999; Andersen, Harrison, Lau, and Rutström 2013; Sutter, Kocher, Glätzle-Rützler, and Trautmann 2013; Halevy 2015). In this literature, little efforts are made to explore the channels of the magnitude effect. This is mainly because most studies employed a single-reward task. With a single-reward task, one cannot disentangle different channels and can only attribute all effects to one aggregate measure, the monetary discount rate.

We are interested in the following question: does the magnitude affect choices through the discount rate, or through intertemporal substitutability (utility curvature)? It is interesting to disentangle channels because different channels predict choices differently in a setting more general than a single-reward task. For instance, in intertemporal allocation tasks with varying interest rates, the discount rate affects the average choices, while the intertemporal substitutability affects how fast a choice changes with the interest rate. The knowledge about mechanisms is important for establishing deeper and better-founded descriptive theories of intertemporal decision making. Meanwhile, omitting a channel of the magnitude effect in an empirical study or in policy making may lead to biased estimates and predictions.

Some existing theories give explanations to the magnitude effect in single-reward tasks. Benhabib, Bisin and Schotter (2010) posit that a fixed cost of waiting makes people impatient to small outcomes, but it matters less when outcomes are large. Noor (2011) establishes a magnitude-dependent discounting model where the discount rate of a dated outcome is decreasing in the size of the outcome. Fudenberg

<sup>2</sup> Frederick, Loewenstein and O'Donoghue (2002, Section 4.2.2) summarized the early literature on the magnitude effect of time preferences. Andersen, Harrison, Lau and Rutström (2011) also reviewed the more recent literature.

and Levine (2006) predict that people exert costly self-control when stakes are high but indulge themselves when stakes are low, which generates a magnitude effect. Holden and Quiggin (2017) assume that people take into account more background consumption when experimental rewards are larger, which also explains the magnitude effect in single-reward tasks. When those theories are applied to intertemporal allocation tasks with proper extension, Benhabib et al. (2010) and Noor (2011) predict a magnitude effect on the discount rate, while Fudenberg and Levine (2006) and Holden and Quiggin (2017) predict a magnitude effect on the utility curvature. It is thereby interesting to know what kinds of models capture the main characteristics of intertemporal choices in a more general setting.

We employ the Convex Time Budget (CTB) method introduced by Andreoni and Sprenger (2012). It allows subjects to form a portfolio of a sooner reward and a later reward given a budget constraint. This intertemporal allocation task has a few applications in the real life. For instance, people need to decide how to allocate their income to the present and to their years of retirement, given the after-tax real rate of return. More importantly, the possibility for subjects to make interior choices (and not only corner choices as in single-reward tasks) enables researchers to simultaneously identify the discount rate and the intertemporal substitutability.<sup>3, 4</sup>

The design of our experiment has three main features. First, all subjects receive equal amounts of participation fees on the sooner date and on the later date regardless of their choices, and the payment conditions are constant across time. Thus, the transaction costs and the trustworthiness of the payments are equalized

<sup>3</sup> Abdellaoui, Bleichrodt and l'Haridon (2013a) provided another method for measuring intertemporal preferences parametrically. Their method identifies utility curvature from marginal utilities for different quantities on the same date, while the CTB method identifies utility curvature from sensitivities of choices to interest rates. The two methods are equivalent if the true model is with a stationary period utility function and a magnitude-independent discount function, as assumed in our paper. If the condition is not satisfied, the former method is better at measuring curvature of a period utility function, and the CTB method is better at measuring sensitivity of choices to interest rates (or, elasticity of intertemporal substitution).

<sup>4</sup> A discussion about the rationality of subjects in the CTB task is provided in Appendix A.

across periods, and these confounding factors are controlled for. Second, we implement two treatments. In one treatment subjects allocate between today and four weeks later, while in the other treatment subjects allocate between four weeks later and eight weeks later. This allows us to assess whether the magnitude effect is affected by the inclusion of a front-end delay. Finally, by assuming a simple yet popular model, the CTB method allows us to identify the discount rate and the atemporal utility function simultaneously. As a result, we are able to disentangle the channels of the magnitude effect.

We find evidence of the magnitude effect in intertemporal allocation tasks: the budget share allocated to the later date is increasing in the total budget. The size of the magnitude effect is found to be decreasing in the magnitude. The pattern is not affected by whether or not a front-end delay is present. At the aggregate level as well as at the individual level, we find magnitude effects both on the discount rate and on intertemporal substitutability. Both channels have considerable impacts on predicted choices. We find that the latter effect is not the same as the magnitude effect on risk attitudes found in previous studies, and hence it might be problematic to correct for the curvature of utility functions by risk attitudes. Instead, the magnitude effect on intertemporal substitutability is consistent with theories where people integrate experimental rewards with more background wealth as the magnitude of rewards gets larger.

The remaining part of this paper is structured as follows: We introduce our experimental design in Section I. In Section II we formulate our hypotheses. In Section III, we investigate non-parametrically the magnitude effect and its relation with the front-end delay. We explore the channels by parametric estimation both at the aggregate level and at the individual level in Section IV. In Section V, we discuss the interpretations of our findings. We draw conclusions in Section VI.

## I. Experimental Design

### A. The Convex Time Budget Method, Parameters and Implementation

The foundation of our experimental design is the Convex Time Budget method introduced by Andreoni and Sprenger (2012). The method consists of a set of intertemporal allocation tasks: in each *decision* subjects are asked to allocate  $N$  tokens to two dates,  $t$  days from today and  $(t + \tau)$  days from today. Each token allocated to  $t$  is worth  $P_t$  euros, while each token allocated to  $(t + \tau)$  is worth  $P_{t+\tau}$  euros. Suppose a subject allocates  $n_t$  tokens to the sooner date and  $n_{t+\tau}$  to the later date, the amount of the sooner reward will be  $z_t = P_t \cdot n_t$  euros and the amount of the later reward will be  $z_{t+\tau} = P_{t+\tau} \cdot n_{t+\tau}$  euros.

Choices are subject to the budget constraint,  $n_t + n_{t+\tau} \leq N$ , and the non-negativity constraints,  $0 \leq n_t, n_{t+\tau} \leq N$ . Subjects are told that they can allocate any number of tokens they like to one of the two dates. Examples of both corner choices and interior choices are given to remove any hesitation in making either type of choices.

Decisions with the same total budget,  $N$ , are grouped in one *decision form*, which is displayed on one page. There are seven decisions in each decision form. The return to each token allocated to the later date is fixed as  $P_{t+\tau} = \text{€}0.20$ , while the return to each token allocated to the sooner date is varied and takes the values  $P_t = \text{€}0.20, \text{€}0.19, \text{€}0.18, \text{€}0.17, \text{€}0.16, \text{€}0.15$ , and  $\text{€}0.14$ . Hence, those returns imply seven gross interest rates,  $R = 1, 1.05, 1.11, 1.18, 1.25, 1.33$ , and  $1.43$ , respectively, over a period of  $\tau$  days. The constraints can be rewritten as

$$\begin{aligned} R \cdot z_t + z_{t+\tau} &\leq m \\ z_t, z_{t+\tau} &\geq 0 \end{aligned}$$

where  $m$  is the total budget and  $m = P_{t+\tau} \cdot N$ .

We implement the CTB method by a zTree program (Fischbacher 2007). Figure 1 shows the interface of a typical decision form. Each decision takes a row. Decisions can be made by scrolling the bars. Once an adjustment is made for one decision, the amounts of the sooner reward and of the later reward in that decision are automatically calculated and displayed.

To avoid any possible effects of initial values, the amounts of rewards are initially blank. Decisions cannot be submitted until all the scrollbars have been adjusted at least once.

### *B. Procedures*

There are two parts in our experiment. Part I consists of five decision forms, with  $N = 100, 200, 300, 400,$  and  $800$ . The order is randomly drawn for each subject. Subjects can move to a specific decision form by clicking the button with the corresponding number. One can go to any decision form at any time, regardless of whether the current decision form is completed. Decisions are automatically stored when one switches to another decision form. This makes comparisons across magnitudes very easy to the subjects in case they would want to make such comparisons. Decisions can only be submitted when all the 35 decisions in the five decision forms are completed.

We randomly assign subjects to one of two treatment groups. In the Present Group, the sooner date is today while the later date is four weeks from today, i.e.  $t = 0$  and  $\tau = 28$ . In the Delayed Group, the sooner date is four weeks from today while the later date is eight weeks from today, i.e.  $t = 28$  and  $\tau = 28$ . Comparing the two groups enables us to check if there exists a present bias on average, and more importantly, if there exists a magnitude effect when no rewards are available in the present.



Form 1: Decisions 1 - 7

September 2014	October 2014	November 2014	December 2014
1 2 3 4 5 6 7	1 2 3 4 5	1 2	1 2 3 4 5 6 7
8 9 10 11 12 13 14	6 7 8 9 10 11 12	3 4 5 6 7 8 9	8 9 10 11 12 13 14
15 16 17 18 19 20 21	13 14 15 16 17 18 19	10 11 12 13 14 15 16	15 16 17 18 19 20 21
22 23 24 25 26 27 28	20 21 22 23 24 25 26	17 18 19 20 21 22 23	22 23 24 25 26 27 28
29 30	27 28 29 30 31	24 25 26 27 28 29 30	29 30 31

You are required to allocate 300 tokens between the two dates for each decision:

	Payment A: 20 Oct 4 WEEKS from Today	and	Payment B: 17 Nov 8 WEEKS from Today	20 Oct	17 Nov
1	265 tokens at € 0.20 each on 20 Oct	and	35 tokens at € 0.20 each on 17 Nov	€ 53.00	€ 7.00
2	_____ tokens at € 0.19 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€
3	_____ tokens at € 0.18 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€
4	_____ tokens at € 0.17 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€
5	_____ tokens at € 0.16 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€
6	_____ tokens at € 0.15 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€
7	_____ tokens at € 0.14 each on 20 Oct	and	_____ tokens at € 0.20 each on 17 Nov	€	€

Switch to Form

1	2	3	4	5
---	---	---	---	---

**Submit Decisions**

--Clicking this button will submit ALL your decisions in every form

Figure 1. Interface of a Typical Decision Form in Part I

Part II is composed of an extended CTB decision form with seven decisions. Subjects are asked to allocate 400 tokens to three dates, today, four weeks from today and eight weeks from today. One additional restriction is imposed, depending on which group one is in. A subject in the Present Group can allocate either 0 or 200 tokens to eight weeks from today; she cannot choose other numbers. But she is still free to allocate any number of tokens between today and four weeks from today. Similarly, a subject in the Delayed Group can allocate either 200 or 400 tokens to today. She is still free to allocate any number of tokens (if there remains some) between four weeks from today and eight weeks from today. The restrictions and the returns to one token allocated are shown in Table 1.

Table 1. Restrictions on the number of tokens and returns to one token allocated to a specific date in Part II

Group		Today	Four weeks from today	Eight weeks from today
Present	Returns to one token	€0.20, €0.19, €0.18, €0.17, €0.16, €0.15, €0.14	€0.20	€0.26
	Restriction on the number of tokens	No additional restriction	No additional restriction	0 or 200
Delayed	Returns to one token	€0.08	€0.20, €0.19, €0.18, €0.17, €0.16, €0.15, €0.14	€0.20
	Restriction on the number of tokens	200 or 400	No additional restriction	No additional restriction

The additional date (eight weeks from today for the Present Group or today for the Delayed Group) is accompanied with a very high return for the Present Group and a very low return for the Delayed Group, so that subjects are induced to allocate 200 tokens to this additional date. If they do so, the remaining task is equivalent to the one with a total budget of 200 tokens in Part I. This characteristic makes the two decision forms comparable.

The purpose of Part II is to test the time separability of intertemporal preferences. One alternative hypothesis is that a subject in the Delayed Group may allocate less to the sooner date if she has allocated a large amount of money to an even sooner date, since the desire for extra consumption has already been partly satisfied. A similar hypothesis applies to the Present Group: a subject in the Present Group may allocate less to the later date if she has already allocated a large amount of money to an even later date, since the guilt for not saving has been partly released. If preferences are time non-separable, the use of a model with a time separable preference is more likely to be problematic. Thus, we want to test the hypothesis of time separability before we perform parametric estimation with a time-separable model.

We do not directly give a fixed reward on the additional date. This is because a fixed reward might be mentally isolated from the allocation task due to narrow bracketing, and hence the test of time separability in the allocation task may be invalid.

At the end of the experiment, subjects were asked to finish a questionnaire. As in previous studies with the CTB method, we asked about subjects' expenditures in a typical week. The average response was €55.22 per week or €7.89 per day.

### *C. Experimental Payments*

The payments are composed of two parts. First, all subjects receive a €5 *participation fee* on each of the two dates scheduled in Part I. Second, each subject has a 10% chance to receive *earnings from decisions*. Before the experiment starts, each subject is randomly given a lottery number, ranging from 0 to 9. After all subjects in a session finish the questionnaire, the experimenter invites one of the subjects to draw a ten-sided die in front of all subjects in the session. Subjects who have a lottery number that equals the die roll get the earnings from decisions. One

decision is randomly selected from the 42 decisions in the two parts as the *decision that counts*. If the decision that counts is from Part I, the allocation in that decision will be realized as the earnings from decisions. If the decision that counts is from Part II, the allocation will be realized and the subject will also receive a €5 participation fee on the additional date in Part II; hence a subject will receive three participation fees if a decision in Part II is realized. All the rules above were articulated in the instructions, and the instructions were always read aloud before either part of the experiment.

The earnings were paid by bank transfer to subjects' checking accounts. We made orders of transfers soon after the experiment and sent reminder emails with information about the incoming amounts on the experimental day and on all the payment dates. Given the reliability of the banking service, subjects can expect to receive all delayed payments exactly on the appropriate payment dates, while some of the present payments might be received one day after the experimental day due to the inter-bank processing.

We believe the payment tool we used was as good as cash in terms of liquidity. Checking accounts are used in private transactions such as paying for rents. Checking accounts are also linked to debit cards. In the Netherlands, debit cards are widely used for daily transactions in almost all kinds of stores including supermarkets, university restaurants and bookstores without any transaction fees. We held a survey about subjects' use of debit cards in the questionnaire. The responses show that bank transfers give high liquidity to the rewards, so that no isolation effect should be expected due to the payment method.<sup>5</sup>

<sup>5</sup> 84.7% of the subjects pay at least 50% of their expenditure in general by debit card, while 91.1% pay at least 30% of their expenditure in general by debit card. Among those who pay less than 30% of their expenditure in general by debit card, 61.1% pay at least 30% of their expenditure in university restaurants or in supermarkets by debit card. Among the remaining seven subjects, four withdraw cash at least 3 times per month.

#### *D. Transaction Costs and Credibility of Payments*

For our experiment, it is extremely important to equalize the transaction costs and the trustworthiness of the payments across periods, because a difference in the transaction costs over the two periods can be a confounding factor of the magnitude effect.

Several facilities were employed in order to equalize the transaction costs across periods and to increase the credibility of the payments. The transaction costs include the costs to collect rewards, to confirm that the rewards have been received with correct amounts, and to remember the earnings so that they can be consumed on the expected dates.

First, we sent reminder emails with information about the incoming amounts on the experimental day and on all the payment dates. Subjects knew this from the instructions, so they did not need to worry about forgetting the earnings on the payment dates, a situation in which the expected marginal utility of the delayed rewards might be lowered.

Second, as Andreoni and Sprenger (2012) did, we delivered our business card and told the subjects to contact us immediately in case they would not receive a payment on time. It increased the credibility of payments and meanwhile served as a reminder of the payments.

Third, we asked subjects to fill in a payment reminder card with the amounts of their rewards on the corresponding dates just after their earnings were displayed. This served as a second reminder in case they forget to check emails.

In sum, the characteristics that one will receive a participation fee on each payment date and that all payments will be received by bank transfer help equalize the transaction costs of receiving payments on all dates. At the same time, the business cards, the payments reminder cards and the reminder emails reduced the risk of forgetting the rewards. The business cards also lowered the perceived default

risks. Even though the risk might still be perceived by some subjects, it should be equal across periods since the payment tools and all auxiliary facilities were the same.

### *E. Sample*

Our experiment was conducted at the CentERlab, Tilburg University in September of 2014.<sup>6</sup> 203 students of the university participated in one of the 11 sessions, 94 in the Present Group and 109 in the Delayed Group. Each subject made 42 decisions. One session took one hour and ten minutes on average. 22 subjects got the earnings from decisions, which averaged €69.16. The overall average earning was €17.49.

## **II. Hypotheses**

Most previous studies define the magnitude effect in single-reward tasks. Denote a reward  $z_t$  on a sooner date  $t$  by  $(z_t, t)$ . In a single-reward task, a subject chooses between a sooner reward  $(z_t, t)$  and a later reward  $(z_{t+\tau}, t + \tau)$ , where  $\tau > 0$  is the delay. A subject displays a *(positive) magnitude effect* if for all  $z_t > 0, z_{t+\tau} > 0$  and  $\eta > 1$ ,

$$(z_t, t) \sim (z_{t+\tau}, t + \tau) \implies (\eta z_t, t) < (\eta z_{t+\tau}, t + \tau).$$

In words, the later reward is more favorable if the amounts are scaled up.

We adapt the definition of the magnitude effect to the intertemporal allocation task. A subject makes a choice  $z^*(R, m)$  out of a linear budget set  $\{(z_t, z_{t+\tau}): R \cdot z_t + z_{t+\tau} = m\}$ , where  $z^* = (z_t^*, z_{t+\tau}^*)$ . She displays a *(positive) magnitude effect* if for all  $m > 0$  and  $\eta > 1$ ,

<sup>6</sup>The payment dates were in September, October and November. The fall semester in Tilburg University started from the end of August and ended in early December. Hence the payment dates were earlier than the final exam weeks and the Christmas vacation, which keeps our experiment from their probably large impacts on the subjects' demand of money.

$$\frac{z_{t+\tau}^*(R, m)}{m} \leq \frac{z_{t+\tau}^*(R, \eta m)}{\eta m}$$

and

$$z_t^*(R, m) > 0 \Rightarrow \frac{z_{t+\tau}^*(R, m)}{m} < \frac{z_{t+\tau}^*(R, \eta m)}{\eta m}.$$

In words, people like to put a larger share of budget on the later date as the total budget is increased. She may fail to do so because the sooner reward is already zero. The adapted definition is consistent with the original one, because in both cases, options with a larger fraction of later reward become more favorable if all options in the menu are scaled up in amount.

*Hypothesis 1 (magnitude effect on budget share):  $\frac{z_{t+\tau}^*}{m}$  is increasing in  $m$ .*

We are also interested in whether the magnitude effect is affected by the presence of a front-end delay. Benhabib et al. (2010) suggest that a fixed cost of delaying rewards can account for the magnitude effect in single-reward tasks, since the fixed cost matters less as the rewards are scaled up in amount. However, it is not clear if this cost is incurred only when a present reward is delayed or if it applies equally to delaying a future reward. We thus test whether the magnitude effect is smaller or even non-existent if the sooner reward is also in the future.

*Hypothesis 2 (a front-end delay leads to a smaller magnitude effect):  $\frac{z_{t+\tau}^*}{m}$  changes less with  $m$  in the Delayed Group than in the Present Group.*

The two hypotheses above can be tested without assuming a specific model.

Conditional on finding a positive magnitude effect, we wish to explore the channels of the magnitude effect. As long as time separability is supported by our results, we will estimate the parameters of preferences, with the assumption that subjects maximize a time separable utility function with CRRA atemporal utility functions and quasi-hyperbolic discounting, i.e. subjects maximize

$$(1) \quad U(z_t, z_{t+\tau}) = \frac{1}{\alpha} \delta^t (z_t + \omega)^\alpha + \beta \delta^{t+\tau} \frac{1}{\alpha} (z_{t+\tau} + \omega)^\alpha,$$

where  $\beta$  is the present bias parameter,  $\delta$  is the daily discount factor,  $\alpha$  is the exponent parameter.  $z_t$  and  $z_{t+\tau}$  are the sooner reward and the later reward, respectively.  $\omega$  is the background consumption mentally integrated with the experimental reward when the decision is made.

When the CRRA utility function is assumed, the elasticity of intertemporal substitution in consumption,  $e_c \equiv -\frac{\ln\left(\frac{c_{t+\tau}}{c_t}\right)}{\ln\left(\frac{u'(c_{t+\tau})}{u'(c_t)}\right)}$ , is equal to  $\frac{1}{1-\alpha}$  ( $c_t$  and  $c_{t+\tau}$  are the consumption on the sooner date and on the later date, respectively.). Thus, the exponent parameter,  $\alpha$ , is a positive transformation of  $e_c$ . If  $\alpha \rightarrow 1$ , the atemporal utility function becomes linear, and the elasticity goes to infinity. In that case, subjects just go for the largest present value, and hence rewards are perfectly substitutable between dates. In case  $\alpha \rightarrow -\infty$ , the atemporal utility function is Leontief, and the elasticity goes to zero. In that case, subjects always divide the total budget into two equal amounts. In general, the larger the value of  $\alpha$ , the more substitutable the subject considers the two rewards to be. Therefore,  $\alpha$  is a measure of intertemporal substitutability.

It brings several advantages to assume such a model. First, the parameters in this model have important economic meanings. The discount factor determines the average choice across interest rates and hence measures the patience of the subject; if a subject is more patient, she will allocate more tokens to the later date for all interest rates. The intertemporal substitutability of consumption between different points in time relates to the dispersion of the choices across interest rates since it measures how sensitive the choices are to the interest rate. These behavioral measures are hard to estimate without assuming a model. Due to the non-negativity constraint, choices are censored at the corners if the preference parameters are extreme. As a result, directly measuring the average choice (as a measure of



patience) and the dispersion of choices (as a measure of intertemporal substitutability) leads to biases. In contrast, the model we assume is tractable and easy to estimate. Moreover, the model is widely used in both theoretical and empirical applications.<sup>7</sup>

Given the model above, we test the following two hypotheses.

*Hypothesis 3 (magnitude effect on discount factor):  $\delta$  is increasing in  $m$ .*

*Hypothesis 4 (magnitude effect on intertemporal substitutability):  $\alpha$  is increasing in  $m$ .*

<sup>7</sup> To address the concern about misspecification, in Appendix C, we check the robustness of our results by estimating a model with the Hyperbolic Absolute Risk Aversion (HARA) utility function and quasi-hyperbolic discounting. The HARA utility function is more flexible in the sense that it allows the atemporal utility function to be Increasing, Constant or Decreasing Relative Risk Aversion. This kind of flexibility is especially important when the magnitude is varied in the experiment. The results are the same.

### III. Overall Effects

#### A. Magnitude Effect on Budget Share

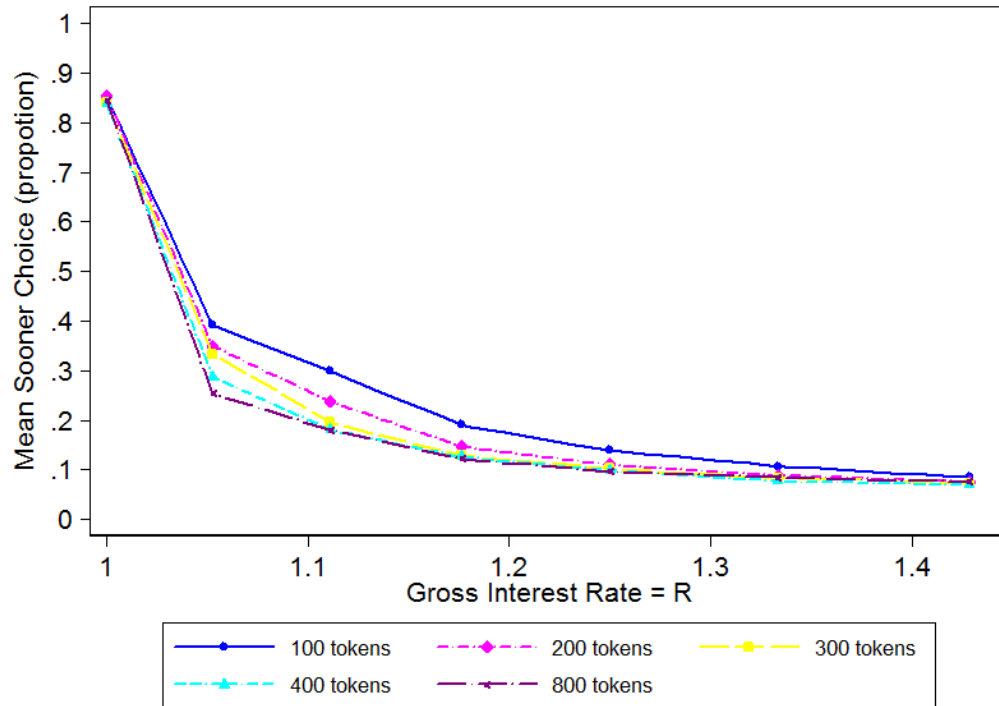


Figure 2. Mean Budget Share on the Sooner Date in Part I

In Figure 2 we plot the mean budget share allocated to the sooner date against the gross interest rate,  $R$ , of each CTB decision in Part I.<sup>8</sup> We plot separate points for the five magnitudes ( $m = \text{€}20, \text{€}40, \text{€}60, \text{€}80, \text{€}160$ ). The budget share allocated to the sooner date declines with the magnitude.

<sup>8</sup> In our data, 28% of the choices are interior, and 62% of our subjects make at least one interior choice. This is very close to the 30% and 63%, respectively, in Andreoni and Sprenger (2012). The relationships between the budget shares and the interest rates are also similar.

The difference seems to be larger when the interest rate is smaller but still positive. This is mainly due to censoring. When the interest rate is zero ( $R = 1$ ) or the highest ( $R = 1.43$ ), most choices are at the corners for both smaller and larger magnitudes.

To judge whether there is a significant magnitude effect, we perform Hotelling's T-squared tests on the mean differences in budget shares between magnitudes, taking seven choices with the same magnitude as a vector (see Table 2).<sup>9</sup> The null hypothesis is that the means of choices are the same across magnitudes, taking into account the correlation within subject. This class of tests makes sense because individual heterogeneity may have made different subjects reveal magnitude effects on tasks with different interest rates (e.g. Subject 1 on Interest Rate 1 while Subject 2 on Interest Rate 2), so that the magnitude effects on all choices would be jointly significant, but the effect on choices with any single interest rate might be insignificant. The results show that the magnitude effect is significant between the magnitudes of €20 and €40 and between any two non-adjacent magnitudes. These results support Hypothesis 1, which states that a larger share of the budget is allocated to the later date when the size of the budget increases.<sup>10</sup>

The results also show that the differences are insignificant between adjacent magnitudes larger than €20. Since the allocation is monotonic in the magnitude and the differences are significant between non-adjacent magnitudes, the insignificance suggests that the magnitude effect is the greatest when comparing the smallest magnitudes (€20 and €40), and becomes smaller for larger magnitudes. The pattern is consistent with the fact that Andersen et al. (2013) found a “statistically

<sup>9</sup> Hotelling's T-squared test is asymptotically nonparametric, so it can be applied to a large sample in nonnormal cases. We also perform a multivariate signed-rank test (Oja and Randles 2004) and the results are basically the same: the magnitude effects are significant between the magnitudes of €20 and €40 and between any two non-adjacent magnitudes at least at the 10% level.

<sup>10</sup> In Table 2 statistics are reported only for four pairs of non-adjacent magnitudes. The mean differences for the other three pairs of non-adjacent magnitudes (in the two groups separately and in total) are also significant: seven out of the nine differences are significant at the 1% level, while the other two at the 5% level.

significant” but “not economically significant” magnitude effect when they elicited time preferences using very high stakes.<sup>11</sup>

Table 2. Multivariate mean difference tests between magnitudes

		€20 ~ €40	€40 ~ €60	€60 ~ €80	€80 ~ €160	€20 ~ €60	€40 ~ €80	€60 ~ €160
Total	F-statistic	3.3184***	1.5477	1.5761	1.7104	5.1162***	3.9336***	3.0328***
	p-value	0.0023	0.1533	0.1444	0.1084	0.0000	0.0005	0.0047
Present Group	F-statistic	2.4091**	1.2374	1.6696	1.3919	3.2252***	2.9725***	2.0495*
	p-value	0.0266	0.2913	0.1270	0.2190	0.0044	0.0076	0.0577
Delayed Group	F-statistic	2.4388**	1.3650	1.0096	1.3659	2.8009**	2.1048**	2.0595*
	p-value	0.0237	0.2282	0.4290	0.2278	0.0104	0.0495	0.0547

*Notes:* Hotelling’s T-squared tests on the mean differences in the budget shares between two magnitudes for all gross interest rates. 203 sets of observations for each magnitude. The degrees of freedom of the F-statistics are (7,196) in total, (7, 87) in the Present Group and (7, 102) in the Delayed Group. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

The results are robust against multiple hypotheses testing problem, since performing a Holm-Bonferroni correction on any family of four/three hypotheses does not change the significance at the 10 percent level.

### *B. Conditional on the Presence of an Immediate Reward?*

We test if the magnitude effect is affected by the existence of a frond-end delay. First, Table 2 shows the results of the Hotelling’s T-squared tests on the magnitude effects for the Present Group and for the Delayed Group, respectively. We find significant magnitude effects in both groups. This implies that the presence of an immediate reward is not a necessary condition for the magnitude effect.

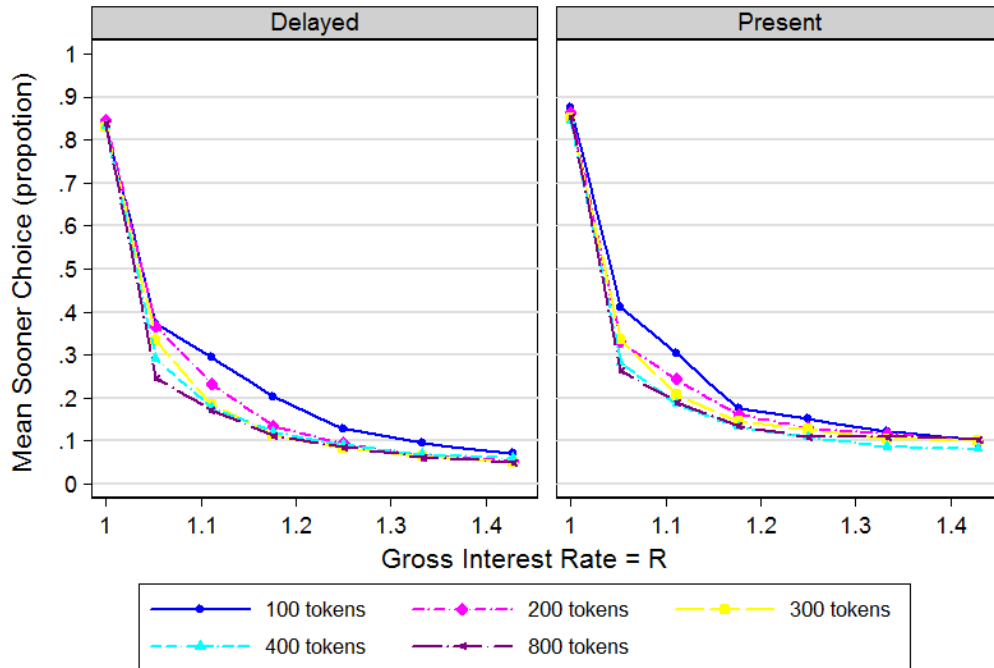
Second, we plot separate graphs for the two groups in Figure 3. Subjects in the Delayed Group seem to be slightly more patient than those in the Present Group. However, when we perform the Hotelling’s T-squared test on all the 35 decisions

<sup>11</sup> Andersen et al. (2013) compare magnitudes of 1,500 Danish kroner and 3,000 Danish kroner, which are roughly equivalent to €200 and €400, respectively.

in Part I between groups, the null hypothesis that the two groups have the same mean responses is not rejected. The p-value is 0.2424 when the degree of freedom is (35, 167). Thus, we find no evidence that the magnitude effect is affected by the existence of a front-end delay, and we reject Hypothesis 2.<sup>12</sup>

This finding has an implication to the modeling of the magnitude effect: no matter what generates the magnitude effect, it applies equally to situations with an immediate reward and those without. For instance, if it is a fixed cost of delaying rewards that generates the magnitude effect, as proposed by Benhabib et al. (2010), the cost applies equally to delaying an immediate reward and to delaying a future reward.

<sup>12</sup> This result also implies that a present bias is not detected. The present bias here refers to non-stationarity of preferences according to the categorization of Halevy (2015). The result of no present bias is consistent with Sutter et al. (2013), and does not necessarily contradict the stylized fact that the discount rate is decreasing in the time distance between the sooner reward and the later reward, as in Benhabib et al. (2010). That stylized fact and stationarity can hold simultaneously if there is subadditivity in discounting (see Read 2001).



Graphs by Group

Figure 3. Mean Budget Shares in the Present Group and in the Delayed Group

### C. Time Separability

Our results show that Part II is a valid test of time separability, since most subjects chose 200 tokens for the additional date in Part II. Only nine out of 203 subjects selected a different number than 200 to the additional date, which involved 41 (2.9%) out of 1415 decisions.

After removing those decisions, we compare the choices with the magnitude of €40 between Part I and Part II, separately for each group. Table 3 shows that the Hotelling's T-squared tests fail to reject the null hypothesis that responses to the two parts have the same means. Those results support time separability, which will be assumed in the next section.

Table 3. Multivariate mean difference tests between Part I and Part II

Subsample	Present	Delayed	Total
F-statistic	1.5560	1.4192	1.0979
Degree of freedom	7, 79	7, 101	7, 187
p-value	0.1609	0.2058	0.3662

*Notes:* Hotelling’s T-squared tests on the mean differences in the budget shares in the decisions with the magnitude of €40 between Part I and Part II. Subjects who chose a different number from 200 tokens for the additional date in Part II such that their choices were not comparable between the two parts have been removed from the sample. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

## IV. Channels

In order to disentangle the magnitude effect into two channels, we perform parametric estimations both at the aggregate level and at the individual level. We then test if the preference parameters change with the magnitude of the total budget.

### A. Aggregate-Level Estimation

#### 1. Estimation strategy

In our main specification, we assume a time-separable utility function with CRRA atemporal utility functions as in equation ( 1 ). We set  $\omega$  (background consumption) equal to the average response to the question about one’s typical daily expenditure, €7.89, as Andreoni and Sprenger (2012) did in two of their specifications.<sup>13</sup>

Given the intertemporal utility function, solving the optimization problem yields the tangency condition

<sup>13</sup> To fix the background consumption across subjects brings the advantage that all effects come from the variation in choices rather than also from the variation in the self-reported background consumptions, which may be noisy. We check the robustness by setting  $\omega$  as individual background consumption, and average/individual background consumption combined with the participation fee (See Appendix B). The results are basically the same.

$$\frac{z_t + \omega}{z_{t+\tau} + \omega} = \begin{cases} (\beta \delta^\tau R)^{\frac{1}{\alpha-1}}, & \text{if } t = 0 \\ (\delta^\tau R)^{\frac{1}{\alpha-1}}, & \text{if } t > 0 \end{cases}.$$

Taking logs gives a linear equation

$$\ln\left(\frac{z_t + \omega}{z_{t+\tau} + \omega}\right) = \left(\frac{\ln \beta}{\alpha - 1}\right) \cdot 1_{t=0} + \left(\frac{\ln \delta^\tau}{\alpha - 1}\right) + \left(\frac{1}{\alpha - 1}\right) \cdot \ln R$$

where  $1_{t=0}$  is the indicator for the Present Group.

The parameters to be estimated are the present bias parameter,  $\beta$ , the discount factor,  $\delta$ , and the CRRA curvature parameter,  $\alpha$ . The present bias parameter is identified by the differences in allocation between the Present Group and the Delayed Group. If there is a present bias, subjects in the Present Group will allocate more tokens to the sooner date than those in the Delayed Group. The discount factor is identified by one's average choice across different experimental interest rates. A more patient subject will allocate more tokens to the later date in all decisions. The curvature parameter is identified by the dispersion of one's choices across interest rates. Those who consider rewards highly substitutable over time are likely to make corner choices in all decisions, while those with lower elasticity of intertemporal substitution will make choices closer to equal splits.

Following the practice in previous studies (Andreoni and Sprenger 2012; and Augenblick, Niederle and Sprenger 2015), we assume a normally distributed error term additive to the log allocation ratio and take censoring into consideration, then we yield the two-limit Tobit model:

$$\begin{aligned} l_{i,j,k}^* &\equiv \ln\left(\frac{z_{t;i,j,k}^* + \omega}{z_{t+\tau;i,j,k}^* + \omega}\right) \\ &= \left(\frac{\ln \beta}{\alpha - 1}\right) \cdot 1_{t=0;i} + \left(\frac{\ln \delta^\tau}{\alpha - 1}\right) + \left(\frac{1}{\alpha - 1}\right) \ln R_j + \epsilon_{i,j,k}, \epsilon_{i,j,k} \sim N(0, \sigma_k) \end{aligned}$$



$$l_{i,j,k} = \begin{cases} \ln \frac{\omega}{m_k + \omega}, & \text{if } l_{i,j,k}^* \leq \ln \frac{\omega}{m_k + \omega} \\ l_{i,j,k}^*, & \text{if } \ln \frac{\omega}{m_k + \omega} < l_{i,j,k}^* < \ln \frac{\frac{m_k}{R_j} + \omega}{\omega} \\ \ln \frac{\frac{m_k}{R_j} + \omega}{\omega}, & \text{if } l_{i,j,k}^* \geq \ln \frac{\frac{m_k}{R_j} + \omega}{\omega} \end{cases}$$

where  $i = 1, \dots, 203$  denotes Subject  $i$ ,  $j = 1, \dots, 7$  denotes Interest rate  $j$ , and  $k = 1, \dots, 5$  denotes Magnitude  $k$ . The error term is allowed to vary across magnitudes since giving a larger number of tokens might induce a larger noise, which might be a competing explanation of a larger sensitivity to the interest rate.

The model is estimated by the quasi-maximum-likelihood method: when performing the estimation, the error term,  $\epsilon$ , is assumed to be i.i.d., while in computing the standard errors, the error term is assumed to be independent across subjects, but might be correlated within-subject. Estimates of the parameters can be recovered and standard errors can be inferred by the delta method.

Since we are interested in the magnitude effect, we also perform the estimation with interaction terms of the parameters and the magnitude dummies. Thus, tests can be performed on the differences between the parameters for different magnitudes.

In Appendix C, we assume another specification, in which the utility function is Hyperbolic Absolute Risk Aversion (HARA). In that specification the background consumption,  $\omega$ , is also a parameter to be estimated. In this way, we address the concern that the average self-reported background consumption may not match the true background consumption integrated with the experimental rewards in decision making, or the coefficient of Relative Risk Aversion of the utility function may not be constant (i.e. the CRRA utility function is misspecified). The results are basically the same.

## 2. Results

Table 4 reports the magnitude-invariant estimates and the magnitude-specific estimates of the parameters, respectively. A salient feature is that none of the estimates of  $\beta$  is significantly different from 1, implying no evidence of present bias, which is consistent with our finding in the model-free analysis. The annual discount rate for all magnitudes is 52.7%, which is in the range found by previous studies. The CRRA curvature parameters are always significantly smaller than 1, implying that the subjects on average consider the monetary rewards received on different dates imperfectly substitutable, which is also consistent with other studies (e.g. Andreoni and Sprenger 2012; Andreoni, Kuhn and Sprenger 2015; Cheung 2015; and Augenblick, Niederle and Sprenger 2015).

Most importantly, both the discount factor and the CRRA curvature are increasing in the magnitude. To judge if these magnitude effects are significant: Table 6 presents Wald tests over the differences of parameters between magnitudes.<sup>14</sup> We find significant magnitude effects both on the discount factor,  $\delta$ , and on the exponent parameter,  $\alpha$ , which is a positive transformation of the elasticity of intertemporal substitution. The discount factor is increasing in the magnitude, meaning that the decision weights on later rewards shift upward when the total budget increases. The elasticity of intertemporal substitution is increasing in the magnitude, meaning that the rewards on the two dates are more substitutable to the subjects when a larger total budget is provided. This results in choices closer to the two corners (to which corner depends on whether  $\delta R > 1$ ). Thereby, we verify Hypothesis 3 and Hypothesis 4.

<sup>14</sup> For the other three pairs of non-adjacent magnitudes: the differences in  $\beta$  are not significant, while the differences in  $\delta^\tau$  and in  $\alpha$  are all significant at the 1% level.

Table 4. Discounting and Curvature Parameter Estimates in the Aggregate-level Estimation with the CRRA Specification

Model:	Tobit			Tobit		
Magnitude:	All	€20	€40	€60	€80	€160
Present bias: $\hat{\beta}$	0.989 (0.018)	0.989 (0.021)	0.989 (0.018)	0.986 (0.018)	0.997 (0.017)	0.986 (0.020)
Discount factor over four weeks: $\hat{\delta}^{\tau}$	0.968 (0.012)	0.948 (0.016)	0.961 (0.013)	0.971 (0.012)	0.972 (0.011)	0.982 (0.012)
CRRA curvature: $\hat{\alpha}$	0.955 (0.004)	0.928 (0.007)	0.947 (0.005)	0.952 (0.005)	0.958 (0.004)	0.968 (0.003)
S.e. of the error term: $\hat{\sigma}$	3.699 (0.281)	2.294 (0.200)	2.986 (0.245)	3.369 (0.269)	3.857 (0.307)	5.314 (0.454)
Log-likelihood	-13678.51			-13538.56		
Observations	7,105			7,105		
Uncensored	1,969			1,969		
Clusters	203			203		

*Notes:* Two-limit Tobit estimators. CRRA estimation with  $\omega = 7.89$  (average reported background consumption). Column 1: assuming that parameters are invariant to magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Log-likelihood has been corrected for the transformation of dependent variables. Standard errors calculated via the delta method.

Table 5. Estimates of Parameter Differences between Magnitudes in the CRRA Specification

Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
Present bias: $\hat{\beta}$	-0.000 (0.008)	-0.003 (0.006)	0.011* (0.006)	-0.011 (0.009)	-0.003 (0.010)	0.008 (0.008)	0.000 (0.009)
Discount factor over four weeks: $\hat{\delta}^{\tau}$	0.014** (0.006)	0.010* (0.005)	0.001 (0.005)	0.010** (0.004)	0.024*** (0.008)	0.011* (0.006)	0.011* (0.006)
CRRA curvature: $\hat{\alpha}$	0.018*** (0.003)	0.006*** (0.002)	0.006*** (0.002)	0.011*** (0.002)	0.024*** (0.004)	0.011*** (0.002)	0.016*** (0.002)
S.e. of the error term: $\hat{\sigma}$	0.692*** (0.106)	0.383*** (0.092)	0.487*** (0.108)	1.457*** (0.218)	1.075*** (0.138)	0.871*** (0.125)	1.944*** (0.258)

*Notes:* Estimates of parameter differences are inferred from the Two-limit Tobit estimation by the delta method. The estimation assumes CRRA utility with  $\omega = 7.89$ . Separate parameters are estimated for each magnitude among €20, €40, €60, €80 and €160. There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

To get an idea about the relative importance of the two channels of the magnitude effect, we use the estimates above to predict choices in the 35 questions for both the Present Group and the Delayed Group. Table 6 presents the marginal effects of allowing one parameter to vary with the magnitude: in each row, we allow only one parameter, either  $\delta$  or  $\alpha$ , to vary with the magnitude of the decisions (as indicated by the column title), but fix the other two parameters at the value estimated from the magnitude of €20. Each number in a cell is the total change (in unit of  $\frac{N_k}{100}$ , the percentage of the total budget) in the seven decisions with the corresponding magnitude. The results show that the marginal effect of allowing  $\alpha$  to vary with the magnitude is at least as large as the marginal effect of allowing  $\delta$  to vary. This suggests that the magnitude effect on the elasticity of intertemporal substitution is at least as important as the magnitude effect on the discount rate.

Table 6. Marginal Effects of Allowing a Parameter to Vary with Magnitudes in the CRRA Specification

Magnitude:		€40	€60	€80	€160
Parameter values used in prediction:	$\beta_1, \delta_k, \alpha_1$ (Delayed):	21.3	34.7	36.6	48.6
	$\beta_1, \delta_k, \alpha_1$ (Present):	21.8	36.5	38.1	50.7
	$\beta_1, \delta_1, \alpha_k$ (Delayed):	24.8	33.9	44.4	65.5
	$\beta_1, \delta_1, \alpha_k$ (Present):	27.7	38.0	49.2	72.8

*Notes:* The changes in choices predicted by the CRRA Tobit model using the parameter values indicated by the row title compared with  $(\beta_1, \delta_1, \alpha_1)$ , for the two groups separately.  $k$  in the row titles stands for the magnitude in the column title. For instance, the first cell in the first row is the difference between the choices made in the seven decisions with the magnitude of €40 predicted by the model with parameter values  $(\beta_1, \delta_2, \alpha_1)$  and those predicted by the model with parameter values  $(\beta_1, \delta_1, \alpha_1)$ . In other words, it is the marginal effect of allowing  $\delta$  to vary with the magnitude from €20 to €40. The unit is 1 percent of the total budget.

### B. Individual-Level Estimation

The aggregate-level estimation provides evidence of positive magnitude effects on the discount factor and on intertemporal substitutability. One may wonder whether these results also hold at the individual level. Indeed, we find a huge individual heterogeneity in choices. One concern is that, when testing the magnitude effect on the aggregate preferences, there might be a bias resulting from forcing all subjects to have the same preferences and the same distribution of noise. To deal with this concern, we also perform individual-level estimation and tests.

#### 1. Estimation and testing procedure

We keep all the assumptions that underlie equation ( 1 ) except for  $\beta$  since it is not identified in individual-level estimations. We estimate the discount factor ( $\delta$ ) and the intertemporal substitutability ( $\alpha$ ) for each combination of subject and stake, and then test if the two parameters are increasing in the magnitude within-subject.

One important difference from the aggregate-level estimation is that there might be an under-identification problem when a subject made no or only one interior

choice under a stake. Actually, there are 627 out of 1015 (62%) combinations of subjects and stakes suffering from such a problem. We thereby adopt a conservative way to test the magnitude effect. First, we yield point estimates of  $\delta$  and  $\alpha$  if possible. Whenever there is an under-identification problem, we remove the error term from ( 1 ) and then infer the intervals of  $\delta$  and  $\alpha$  that can generate the observations. Second, we perform a one-tailed sign test on the two parameters, respectively, with the null hypotheses that they do not change with the magnitude. The sign test only requires that the distribution of a parameter does not differ between magnitudes, while it allows the distribution to be different across subjects. For a comparison between a point estimate and an interval estimate, we recognize a difference only if the point is not in the interior of the interval. For a comparison between two interval estimates, we recognize a difference if the two intervals do not overlap.

## 2. Results

Table 7 shows the results of the tests at the individual level. We reject the null hypotheses of no magnitude effect on the two parameters, in favor of positive magnitude effects. This shows that the two channels of the magnitude effect on intertemporal choices are robust against individual heterogeneity.

Table 7. Sign tests on preference parameters between magnitudes

		€20 ~ €40	€40 ~ €60	€60 ~ €80	€80 ~ €160	€20 ~ €60	€40 ~ €80	€60 ~ €160
Discount factor over four weeks: $\widehat{\delta^\tau}$	#increase	31%	31%	26%	23%	36%	34%	29%
	/unchanged	/48%	/51%	/56%	/58%	/48%	/50%	/54%
	/decrease	/21%	/19%	/18%	/19%	/16%	/16%	/17%
	p-value	0.0255**	0.0107**	0.0687*	0.2253	0.0001***	0.0003***	0.0086***
CRRRA curvature: $\widehat{\alpha}$	#increase	30%	28%	24%	27%	31%	29%	29%
	/unchanged	/61%	/60%	/63%	/64%	/62%	/61%	/63%
	/decrease	/ 9%	/12%	/13%	/ 9%	/ 8%	/10%	/ 9%
	p-value	0.0000***	0.0002***	0.0101**	0.0000***	0.0000***	0.0000***	0.0000***

*Notes:* Right-tailed sign tests on the differences of parameters between two magnitudes. 203 sets of observations for each magnitude.  $\omega = 7.888$ . When under-identification occurs, interval estimates are yielded for the two preference parameters. A point estimate and an interval estimate are considered to be different only if the point is not in the interior of the interval. Two interval estimates are considered to be different only if their intersection is empty or a singleton. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

## V. Interpretations

The results above imply that when an average subject faces a larger budget in an intertemporal allocation task, she behaves more patiently, but also she regards rewards to be more substitutable between dates.

### A. Relation with the Magnitude Effect on Risk Aversion

According to the Discounted Expected Utility (DEU) theory, the risk attitude and the elasticity of intertemporal substitution are represented by the same parameters, since risk aversion and imperfect fungibility both originate from diminishing marginal utility. Therefore, one may wonder whether the magnitude effect on intertemporal substitutability is the same as the magnitude effect on risk attitudes.

We find evidence against this equivalence when we compare our results with previous findings. Holt and Laury (2002) investigated the magnitude effect on risk attitudes with Multiple Price List (MPL) questions. They found a significant,

positive magnitude effect: when faced with a larger magnitude, people appear to be more risk averse in terms of the relative risk aversion. This is in the opposite direction against what we find. Their finding suggests an increase in the concavity as the magnitude increases while ours shows a movement towards linearity. This contradiction suggests that the magnitude effect on relative risk aversion is not driving the magnitude effect on intertemporal substitutability.

Some other studies also suggest that risk aversion and intertemporal substitutability should be separated. Andreoni and Sprenger (2012) found no significant correlation at the individual level between the curvature estimated by the CTB method and the risk attitude elicited by the MPL method. Abdellaoui, Bleichrodt, l'Haridon and Paraschiv (2013b), Miao and Zhong (2015) and Cheung (2015) also found that the utility curvature elicited from intertemporal tasks is quantitatively different from that elicited from tasks with risk. We provide evidence from a different perspective: while the previous studies showed that the degrees of concavity are different for the two kinds of utility functions, we show that the degrees of concavity change in opposite directions when the stake is varied.

This finding has implications for both theories and experimental methods. First, it lends support to the theories which separate intertemporal substitutability from risk aversion, such as Epstein and Zin (1989). Second, it casts doubt on the use of a risk-elicitation task to correct for the curvature when eliciting time preferences.

### *B. Relation with Borrowing Constraints*

In theory, a binding borrowing constraint can lead to a magnitude effect on the monetary discount rate in a single-reward task if the background consumption is expected to grow over time, as shown by Epper (2015). However, Meier and Sprenger (2010) found that experimentally elicited long-run discount rates are uncorrelated with credit constraints, suggesting that on average, whether the



borrowing constraint is binding does not affect intertemporal choices in experiments.

Moreover, given the fact that subjects may have savings which provide some but small liquidity, the number of subjects whose borrowing constraints are binding should be increasing in the stake. For this reason, if the borrowing constraint is a main issue, we should observe that the intertemporal substitutability is decreasing in the stake, which is inconsistent with our results. Therefore, we believe that a binding borrowing constraint is not the main driver of our results.

### *C. Relation with Existing Theories*

We discuss the implications of our empirical findings for some theories that may explain the two channels of the magnitude effect.

One model that can account for the magnitude effect on the discount factor was proposed by Benhabib et al. (2010). They developed a model with a fixed cost of delaying rewards. The idea is that whenever a delayed reward is chosen, a fixed cost is incurred, so that as the stake increases, the cost becomes relatively less important and hence the subject appears more patient.

Noor (2011) proposed a model of magnitude-dependent discounting, which leads to similar predictions. In his model, the discount function is increasing in the utility at the later period. As the stake gets larger, the discount function converges to 1.<sup>15</sup>

One theory that can explain the magnitude effect on intertemporal substitutability is an extended version of the dual-self bank-nightclub model of Fudenberg and Levine (2006). In the original model, the agent first chooses the amount of pocket cash when no temptation is present, and then she chooses the amount of consumption when a windfall is available and temptation plays a role. The strategy

<sup>15</sup> Baucells and Heukamp (2012) extend the magnitude-dependent discounting model to allow risks.

for utility maximization is to spend all of the windfall when it is small but try to save some money out of the windfall when it is large. A small windfall is not integrated with the lifelong wealth, because the agent does not bother to perform self-control, but it is worth controlling oneself when the windfall is large. As a result, the utility function for windfalls is much more concave when the size of the windfall is below a certain threshold than when it is above the threshold.

The model can explain a magnitude effect on intertemporal substitutability if we impose the assumption that an agent who anticipates a reward in the future does not immediately adjust her cash allocation plan. Instead, she keeps the anticipated reward in the mental account of windfalls until it is received and part of it is consumed. Only after the remainder is moved into the mental account of savings does she reschedule her future consumption.

When this assumption is used, the model predicts that a subject is likely to make interior choices when the budget is small, i.e., below the threshold induced by the self-control costs. Since the utility function for windfalls is very concave the subject balances extra consumption on the sooner date and on the later date. As the budget increases above the threshold, the subject likes to save part of it for consumption smoothing. Since the utility function for savings is much less concave (close to linear), these savings will be allocated fully to either the sooner date (when the interest rate is small) or the later date (when the interest rate is large). Hence, as the budget increases the intertemporal substitutability increases and it will appear as if the utility function has become less concave (see Appendix D for a simulation).

Another model that can explain the magnitude effect on intertemporal substitutability is the mental zooming theory proposed by Holden and Quiggin (2017). The theory presumes that people integrate more background consumption with the experimental reward as the size of the reward increases. If the budget increases, individuals 'zoom out' as it were, and take a broader perspective in the decision problem. One reason may be that individuals are likely to divide and use

up a bigger windfall over a longer time period. Based on the data collected from their field experiment with Malawian peasants, Holden and Quiggin showed that the magnitude effect on the discount rate in single-reward tasks would disappear if the unobserved background consumption is assumed to be an increasing function of the stake.

In intertemporal allocation tasks, the increasing background consumption can generate a magnitude effect on intertemporal substitutability. To see why, denote the observed elasticity of intertemporal substitution in experimental rewards by  $e_z$ . The relationship between  $e_z$  and preference parameters is

$$e_z = \frac{1}{1 - \alpha} \cdot \frac{\log\left(\frac{z_{t+\tau}}{z_t}\right)}{\log\left(\frac{z_{t+\tau} + \omega}{z_t + \omega}\right)}.$$

Since  $e_z$  is increasing in both  $\alpha$  and  $\omega$ , an increase in  $\alpha$  and an increase in  $\omega$  are competing explanations for the magnitude effect on intertemporal substitutability. If subjects take into account more background consumption as the total budget increases, we would observe a greater sensitivity to the interest rate, i.e. a greater  $e_z$ . When we assume a fixed background consumption, however, the pattern will be attributed to a magnitude effect on  $\alpha$ .

Both the mental-accounting Fudenberg-Levine model and the mental zooming theory point to partial integration with lifelong wealth, which seems to be an important mechanism of the magnitude effect on intertemporal substitutability. Andersen et al. (2018) showed empirically that subjects only partially integrate experimental rewards with wealth in risk preference tasks. While they provide evidence of partial asset integration by exploiting variation in personal wealth, we provide within-subject evidence suggesting that the degree of asset integration is increasing in the stake.

None of the current models can explain both a magnitude effect on the discount factor and a magnitude effect on the intertemporal substitutability. Of course, the

two channels can be explained by a mode-switching model in which individuals are assumed to have different preferences for different stakes. However, a truly unified explanation is still lacking.

## **VI. Conclusion**

Our study investigates the magnitude effect on intertemporal choices in a setting that is more general than a single-reward task, namely the intertemporal allocation task. After adapting the definition to the new task, we verify the existence of a magnitude effect: people allocate a larger share of budget to the later reward as the total budget increases. The magnitude effect is not affected by whether the sooner reward is immediate or in the future. The size of the magnitude effect is smaller when the stakes are higher.

We then look deeper into the effect, by exploring the channels. The results underscore the importance of a dimension which is often overlooked, namely, the intertemporal substitutability. We find evidence that both discount factor and intertemporal substitutability change with the magnitude of rewards. The magnitude effect on intertemporal substitutability is not riven by the magnitude effect on risk attitudes, since the two magnitude effects are in the opposite directions.

Some existing theories may provide explanations for one of the two channels. A cost-of-delay model (Benhabib et al. 2010) or a magnitude-dependent discounting model (Noor 2011; Baucells and Heukamp 2012) can account for a magnitude effect on the discount factor. Models which allow the degree of asset integration (mental accounting) to vary with the size of the budget can explain a magnitude effect on intertemporal substitutability. However, a new theory would be needed to account for both channels simultaneously and in a unified way.

## REFERENCES

- Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2013a). Sign-Dependence in Intertemporal Choice. *Journal of Risk and Uncertainty*, 47(3), pp. 225-253.
- Abdellaoui, M., Bleichrodt, H., l'Haridon, O., & Paraschiv, C. (2013b). Is There One Unifying Concept of Utility? An Experimental Comparison of Utility Under Risk and Utility Over Time. *Management Science*, 59(9), pp. 2153–2169.
- Andersen, S., Cox, J. C., Harrison, G. W., Lau, M., Rutström, E. E., & Sadiraj, V. (2018). Asset Integration and Attitudes toward Risk: Theory and Evidence. *Review of Economics and Statistics*, 100(5), 816-830.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2011). Discounting Behavior and the Magnitude Effect. *Working Paper 2011-05*.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2013). Discounting Behavior and the Magnitude Effect: Evidence from a Field Experiment in Denmark. *Economica*, 80(320), pp. 670-697.
- Andreoni, J., & Sprenger, C. (2012). Estimating Time Preferences from Convex Budgets. *American Economic Review*, 102(7), pp. 3333-3356.
- Andreoni, J., Kuhn, M. A., & Sprenger, C. (2015). Measuring Time Preferences: A Comparison of Experimental Methods. *Journal of Economic Behavior and Organization*, 116, 451-464.
- Augenblick, N., Niederle, M., & Sprenger, C. (2015). Working Over Time: Dynamic Inconsistency in Real Effort Tasks. *Quarterly Journal of Economics*, 130(3), 1067-1115.
- Baucells, M., & Heukamp, F. H. (2012). Probability and Time Trade-Off. *Management Science*, 58(4), pp. 831-842.

- Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-Bias, Quasi-Hyperbolic Discounting, and Fixed Costs. *Games and Economic Behavior*, 69(2), pp. 205-223.
- Chakraborty, A., Calford, E. M., Fenig, G., & Halevy, Y. (2017). External and Internal Consistency of Choices Made in Convex Time Budgets. *Experimental Economics*, 20(3), 687-706.
- Cheung, S. L. (2015). Eliciting Utility Curvature and Time Preference. *Working Paper 2015-01*.
- Epper, T. (2015). Income Expectations, Limited Liquidity, and Anomalies in Intertemporal Choice. *University of St. Gallen, School of Economics and Political Science Discussion Paper No. 2015-19*.
- Epstein, L. G., & Stanley, E. Z. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57(4), pp. 937-969.
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics*, 10(2), 171-178.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, 40(2), pp. 351-401.
- Fudenberg, D., & Levine, D. K. (2006). A Dual-Self Model of Impulse Control. *American Economic Review*, 96(5), pp. 1449-1476.
- Halevy, Y. (2015). Time Consistency: Stationarity and Time Invariance. *Econometrica*, 83(1), pp. 335-352.
- Holcomb, J. H., & Nelson, P. S. (1992). Another Experimental Look at Individual Time Preference. *Rationality and Society*, 4(2), 199-220.
- Holden, S. T., & Quiggin, J. (2017). Bounded Awareness and Anomalies in Intertemporal Choice: Zooming in Google Earth as Both Metaphor and Model. *Journal of Risk and Uncertainty*, 54(1), 15-35.

- Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, 92(5), pp. 1644-1655.
- Kirby, K. N. (1997). Bidding on the Future: Evidence against Normative Discounting of Delayed Rewards. *Journal of Experimental Psychology: General*, 126, pp. 54-70.
- Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin Addicts Have Higher Discount Rates for Delayed Rewards. *Journal of Experimental Psychology: General*, 128, pp. 78-87.
- Meier, S., & Sprenger, C. (2010). Present-Biased Preferences and Credit Card Borrowing. *American Economic Journal: Applied Economics*, 2(1), pp. 193-210.
- Miao, B., & Zhong, S. (2015). Comment on “Risk Preferences Are Not Time Preferences”: Separating Risk and Time Preference. *American Economic Review*, 105(7), pp. 2272-86.
- Noor, J. (2011). Intertemporal Choice and the Magnitude Effect. *Games and Economic Behavior*, 72(1), pp. 255-270.
- Oja, H., & Randles, R. H. (2004). Multivariate Nonparametric Tests. *Statistical Science*, 19(4), pp. 598-605.
- Read, D. (2001). Is Time-Discounting Hyperbolic or Subadditive? *Journal of Risk and Uncertainty*, 23(1), pp. 5-32.
- Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and Uncertainty: Experimental Decisions Predict Adolescents’ Field Behavior. *American Economic Review*, 103(1), pp. 510-531.
- Thaler, R. (1981). Some Empirical Evidence on Dynamic Inconsistency. *Economics Letters*, 8(3), pp. 201-207.

## **Appendix A. Rationality of Subjects in the Convex Time Budget Method**

The CTB method allows subjects to make interior choices, and hence makes it possible to measure discount rates and utility curvature simultaneously. However, Chakraborty, Calford, Fenig and Halevy (2017) found that a proportion of subjects, especially those who make interior choices, violate wealth monotonicity in the CTB dataset of Andreoni and Sprenger (2012) and that the magnitude of wealth monotonicity violations conditional on violating at least once are as large as that generated by uniform random choice, and hence questioned the rationality of subjects in making CTB decisions.

In this appendix, we follow Chakraborty et al. (2017) to examine price monotonicity and wealth monotonicity of our dataset. In specific, we look at fractions of monotonicity violations among all subjects and among subjects who make at least one interior choice, respectively. We also measure the average magnitude of wealth monotonicity violations for those who violate wealth monotonicity at least once, and we compare it with the distribution of the magnitude generated by uniform random choice.

One advantage of our sample in testing wealth monotonicity is that we have 70 pairs of tasks that only differ in total budget for each subject, compared to one pair in Andreoni and Sprenger (2012). This endows our test with greater power than that in Chakraborty et al. (2017).

Table A 1 shows the rationality indices of the full sample and the subsamples as well as those generated by uniform random choice. The fractions of price monotonicity violations (2-3%) are less than the fractions of wealth monotonicity violations (10-20%). A possible reason is that prices are varied within decision forms but total budgets are varied across decision forms, and hence it is easier to make comparison across prices than across total budgets. Nevertheless, those fractions and the magnitudes of violations are much lower than those generated by



uniform random choice, suggesting that our dataset and the subsample of subjects who make interior choices are highly informative.

Table A 1. Rationality of Subjects Compared to Uniform Random Choice

	Fraction of price monotonicity violations	Average magnitude of price monotonicity violations (euros)	Fraction of wealth monotonicity violations	Average magnitude of wealth monotonicity violations (euros)
Full sample	0.022	0.13	0.101	0.82
Subjects who make at least one interior choice	0.037	0.23	0.155	1.03
Subjects who violate wealth monotonicity at least once	0.033	0.26	0.208	1.67
Mean of uniform random choice (Standard deviation)	0.429 (0.006)	8.12 (0.21)	0.297 (0.008)	2.69 (0.10)

*Notes:* The first three rows present fractions of price monotonicity violations and wealth monotonicity violations as well as average magnitudes of violations in terms of euros in the full sample, the subsample of subjects who make at least one interior choice, and the subsample of subjects who violate wealth monotonicity at least once. The last row presents the means and the standard deviations of the same four indices generated by uniform random choice. The means and standard errors are calculated by simulating 10,000 times.

## Appendix B. Parametric Analysis with Different Assumptions on $\omega$

We check the sensitivity of the parameter estimates (Table A 2) and the magnitude effects (Table A 3) to alternative assumptions on the background consumption. The results show that the magnitude effects on the discount factor and on intertemporal substitutability are robust, except when  $\omega$  is assumed to be very small. However, with a small background consumption, one should rarely

make corner choices due to the motivation of consumption smoothing. But we do observe many corner choices in our sample. Hence, the case with a small  $\omega$  is unlikely to be true.

Table A 2. Background Consumption, Parameter Estimates and Likelihood

Model: Magnitude:		Tobit All	€20	€40	Tobit €60	€80	€160
$\omega = 7.89$	Present bias: $\hat{\beta}$	0.989 (0.018)	0.989 (0.021)	0.989 (0.018)	0.986 (0.018)	0.997 (0.017)	0.986 (0.020)
	Discount factor over four weeks: $\hat{\delta}^\tau$	0.968 (0.012)	0.948 (0.016)	0.961 (0.013)	0.971 (0.012)	0.972 (0.011)	0.982 (0.012)
	CRRA curvature: $\hat{\alpha}$	0.955 (0.004)	0.928 (0.007)	0.947 (0.005)	0.952 (0.005)	0.958 (0.004)	0.968 (0.003)
	S.e of the error term: $\hat{\sigma}$	3.699 (0.281)	2.294 (0.200)	2.986 (0.245)	3.369 (0.269)	3.857 (0.307)	5.314 (0.454)
	Log-likelihood	-13678.51		-13538.56			
	<hr/>						
$\omega =$ $\max(\omega_i, 0.01)$	Present bias: $\hat{\beta}$	0.990 (0.018)	0.989 (0.020)	0.990 (0.018)	0.987 (0.015)	0.998 (0.017)	0.987 (0.022)
	Discount factor over four weeks: $\hat{\delta}^\tau$	0.968 (0.012)	0.948 (0.014)	0.962 (0.012)	0.971 (0.011)	0.972 (0.011)	0.982 (0.012)
	CRRA curvature: $\hat{\alpha}$	0.962 (0.004)	0.945 (0.006)	0.956 (0.005)	0.960 (0.004)	0.964 (0.004)	0.972 (0.003)
	S.e of the error term: $\hat{\sigma}$	4.422 (0.358)	3.048 (0.300)	3.716 (0.315)	4.102 (0.328)	4.570 (0.370)	6.115 (0.552)
	Log-likelihood	-13920.57		-13815.64			
	<hr/>						
$\omega = 12.89$	Present bias: $\hat{\beta}$	0.989	0.989	0.989	0.985	0.997	0.986

		(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)
	Discount factor over four weeks:	0.969	0.948	0.962	0.972	0.973	0.982
	$\widehat{\delta}^\tau$	(0.012)	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)
	CRRA curvature: $\widehat{\alpha}$	0.945	0.903	0.932	0.940	0.948	0.963
		(0.005)	(0.010)	(0.007)	(0.006)	(0.005)	(0.004)
	S.e of the error term: $\widehat{\sigma}$	3.013	1.697	2.329	2.698	3.144	4.490
		(0.229)	(0.147)	(0.187)	(0.214)	(0.262)	(0.384)
	Log-likelihood	-13623.90			-13443.15		
$\omega = \omega_i + 5$	Present bias: $\widehat{\beta}$	0.990	0.990	0.990	0.987	0.998	0.987
		(0.018)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
	Discount factor over four weeks:	0.968	0.948	0.962	0.971	0.972	0.982
	$\widehat{\delta}^\tau$	(0.012)	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)
	CRRA curvature: $\widehat{\alpha}$	0.948	0.912	0.936	0.944	0.951	0.964
		(0.005)	(0.009)	(0.006)	(0.006)	(0.005)	(0.004)
	S.e of the error term: $\widehat{\sigma}$	3.213	1.870	2.529	2.899	3.355	4.430
		(0.246)	(0.163)	(0.205)	(0.231)	(0.275)	(0.398)
	Log-likelihood	-13693.32			-13527.03		
$\omega = 0.01$	Present bias: $\widehat{\beta}$	0.990	0.990	0.990	0.987	0.998	0.986
		(0.019)	(0.020)	(0.019)	(0.018)	(0.020)	(0.022)
	Discount factor over four weeks:	0.964	0.945	0.959	0.968	0.969	0.980
	$\widehat{\delta}^\tau$	(0.013)	(0.015)	(0.013)	(0.010)	(0.013)	(0.011)
	CRRA curvature: $\widehat{\alpha}$	0.989	0.989	0.989	0.989	0.989	0.989
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	S.e of the error term: $\widehat{\sigma}$	15.333	14.511	14.500	14.470	15.251	17.613

	(1.216)	(1.239)	(1.154)	(1.115)	(1.220)	(1.478)
Log-likelihood	-15596.78			-15575.13		
Observations	7,105			7,105		
Uncensored	1,969			1,969		
Clusters	203			203		

*Notes:* Two-limit Tobit estimators. Panel 1:  $\omega$  = average reported background consumption. Panel 2:  $\omega$  = individual reported background consumption (except for one subject, we replace the zero consumption with 0.01). Panel 3:  $\omega$  = average reported background consumption plus the participation fee. Panel 4:  $\omega$  = individual background consumption plus the participation fee. Panel 5:  $\omega = 0.01$ . Column 1: assuming that parameters are invariant to magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Log-likelihood has been corrected for the transformation of dependent variables. Standard errors calculated via the delta method.

Table A 3. Background Consumption and Magnitude Effects

	Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
$\omega = 7.89$	Present bias: $\hat{\beta}$	-0.000	-0.003	0.011*	-0.011	-0.003	0.008	0.000
	Discount factor over four weeks: $\delta^T$	0.014**	0.010*	0.001	0.010**	0.024***	0.011*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.018***	0.006***	0.006***	0.011***	0.024***	0.011***	0.016***
	S.e of the error term: $\hat{\sigma}$	0.692***	0.383***	0.487***	1.457***	1.075***	0.871***	1.944***
$\omega = \max(\omega_i, 0.01)$	Present bias: $\hat{\beta}$	0.001	-0.003	0.011*	-0.011	-0.002	0.008	-0.000
	Discount factor over four weeks: $\delta^T$	0.013**	0.010**	0.001	0.010**	0.023***	0.010*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.011***	0.004**	0.004***	0.008***	0.015***	0.008***	0.012***
	S.e of the error term: $\hat{\sigma}$	0.668***	0.386***	0.468***	1.545***	1.054***	0.854***	2.013***
$\omega = 12.89$	Present bias: $\hat{\beta}$	-0.000	-0.003	0.011*	-0.011	-0.003	0.008	0.000
	Discount factor over four weeks: $\delta^T$	0.014***	0.010**	0.001	0.010**	0.024***	0.011*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.028***	0.009***	0.008***	0.014***	0.037***	0.017***	0.022***
	S.e of the error term: $\hat{\sigma}$	0.632***	0.370***	0.446***	1.346***	1.002***	0.815***	1.792***
$\omega = \omega_i + 5$	Present bias: $\hat{\beta}$	0.001	-0.003	0.011*	-0.011	-0.002	0.008	-0.000

	Discount factor over four weeks: $\delta^T$	0.014***	0.010*	0.001	0.010**	0.023***	0.011*	0.011*
	CRRA curvature: $\hat{\alpha}$	0.025***	0.008***	0.007***	0.013***	0.032***	0.015***	0.020***
	S.e of the error term: $\hat{\sigma}$	0.659***	0.371***	0.456***	1.374***	1.030***	0.827***	1.830***
$\omega = 0.01$	Present bias: $\hat{\beta}$	-0.000	-0.002	0.010	-0.011	-0.003	0.008	-0.001
	Discount factor over four weeks: $\delta^T$	0.014**	0.009	0.001	0.011**	0.023***	0.010	0.011*
	CRRA curvature: $\hat{\alpha}$	0.000	-0.000	0.001	0.001***	-0.000	0.000	0.002***
	S.e of the error term: $\hat{\sigma}$	-0.011	-0.030	0.781*	2.362***	-0.041	0.751	3.143***

*Notes:* Estimates of parameter differences are inferred from the Two-limit Tobit estimation by the delta method. Parameters are separately set for each magnitude among €20, €40, €60, €80 and €160. Panel 1:  $\omega$  = average reported background consumption. Panel 2:  $\omega$  = individual reported background consumption (except for one subject, we replace the zero consumption with 0.01). Panel 3:  $\omega$  = average reported background consumption plus the participation fee. Panel 4:  $\omega$  = individual background consumption plus the participation fee. Panel 5:  $\omega = 0.01$ . There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

### Appendix C. Parametric Analysis with Estimation of $\omega$

The CRRA specification with exogenously set  $\omega$  is simple and easy to estimate, however, one may suspect that the average self-reported background consumption does not match the true background consumption integrated with the experimental rewards in decision making, or the utility function is misspecified. In particular, if the utility function is Hyperbolic Absolute Risk Aversion (HARA),  $\omega$  in ( 1 ) does not represent the background consumption but is a mixture of the background consumption and the HARA intercept parameter.

In order to meet the challenges above, we estimate  $\omega$  instead of setting it exogenously. By doing this, we “let the data tell” what values are suitable for  $\omega$ , and we can also identify the magnitude effects on  $\omega$ .

## 1. Model

We assume a normally distributed error term at the choice level. The error term can be arisen either because of idiosyncratic shocks in preference or because of imprecision in placing the scrollbar.<sup>16</sup> In specific, a latent choice is

$$z_{t;i,j,k}^* = \begin{cases} \frac{m - \left( (\beta \delta R_j)^{\frac{1}{1-\alpha}} - 1 \right) \omega}{R + (\beta \delta R_j)^{\frac{1}{1-\alpha}}} + \epsilon_{i,j,k}, & \text{if } t = 0 \\ \frac{m - \left( (\delta R_j)^{\frac{1}{1-\alpha}} - 1 \right) \omega}{R + (\delta R_j)^{\frac{1}{1-\alpha}}} + \epsilon_{i,j,k}, & \text{if } t > 0 \end{cases}, \epsilon_{i,j,k} \sim N\left(0, \frac{\sigma_k}{100} m\right).$$

Then the choices are censored at the two corners so that the observed choices are

$$z_{t;i,j,k} = \begin{cases} 0, & \text{if } z_{t;i,j,k}^* \leq 0 \\ z_{t;i,j,k}^*, & \text{if } 0 < z_{t;i,j,k}^* < \frac{m_k}{R_j} \\ \frac{m_k}{R_j}, & \text{if } z_{t;i,j,k}^* \geq \frac{m_k}{R_j} \end{cases}.$$

This is a two-limit nonlinear censored model, which can be estimated by the quasi-maximum likelihood method.

## 2. Results

Table A 4 reports the estimates of the parameters from the specifications with magnitude-invariant parameters and with magnitude-specific parameters, respectively.

Table A 5 presents the estimates of the parameter differences between magnitudes. We find a significant magnitude effect on the discount rate. The

<sup>16</sup> In the model with a normally distributed error term additive to the log allocation ratio, the estimator of  $\omega$  is nonlinear in the error term. Simulation shows that the estimator of  $\omega$  is severely biased given our sample size, though it is asymptotically consistent. The model we assume here is the same as the one implicitly assumed by Andreoni and Sprenger (2012) when they perform the nonlinear least square (NLS) estimation. The difference is that we employ the quasi-maximum likelihood estimation, by which we take into account censoring.

magnitude effect on the exponent parameter,  $\alpha$ , is only significant between the magnitudes of €20 and €40. This is reasonable since we find a strong magnitude effect on the background consumption and the HARA intercept parameter (i.e.  $\omega$ ), which have explained most of the magnitude effects on intertemporal substitutability.

Table A 4. Discounting and Curvature Parameter Estimates in the Aggregate-level Estimation with the HARA specification

Model:	Nonlinear Censored		Nonlinear Censored			
Magnitude:	All	€20	€40	€60	€80	€160
Present bias: $\hat{\beta}$	1.000 (0.004)	0.994 (0.010)	1.002 (0.005)	1.000 (0.005)	1.000 (0.005)	0.997 (0.005)
Discount factor over four weeks: $\widehat{\delta^F}$	0.953 (0.004)	0.952 (0.008)	0.960 (0.003)	0.962 (0.003)	0.965 (0.004)	0.967 (0.004)
Curvature parameter: $\hat{\alpha}$	0.997 (0.003)	0.964 (0.009)	0.984 (0.006)	0.982 (0.004)	0.984 (0.003)	0.986 (0.002)
Background consumption and HARA intercept: $\hat{\omega}$	9.556 (2.046)	10.713 (2.209)	15.790 (3.364)	22.551 (4.461)	30.834 (5.896)	67.814 (12.598)
S.d. of the error term: $\hat{\sigma}$	63.948 (4.843)	78.491 (6.835)	68.110 (5.468)	62.147 (4.771)	62.426 (5.010)	65.890 (5.700)
Log-likelihood	- 13136.00			-12836.01		
Observations	7,105			7,105		
Uncensored	1,969			1,969		
Clusters	203			203		

*Notes:* Quasi-Maximum Likelihood estimators. Column 1: assuming that parameters are the same across magnitudes. Column 2-6: assuming that parameters vary with magnitudes. Clustered standard errors in parentheses. Standard errors calculated via the delta method.

Table A 5. Estimates of Parameter Differences between Magnitudes in the HARA specification

Magnitude:	€40 - €20	€60 - €40	€80 - €60	€160 - €80	€60 - €20	€80 - €40	€160 - €60
Present bias: $\hat{\beta}$	0.00807 (0.00552)	-0.00180 (0.00206)	0.00038 (0.00241)	-0.00303 (0.00249)	0.00572 (0.00578)	-0.00197 (0.00255)	-0.00265 (0.00296)
Discount factor over four weeks: $\hat{\delta}^{\bar{r}}$	0.00734 (0.00570)	0.00252* (0.00138)	0.00258 (0.00169)	0.00267* (0.00141)	0.00986* (0.00574)	0.00510*** (0.00176)	0.00526** (0.00215)
Curvature parameter: $\hat{\alpha}$	0.02007*** (0.00737)	0.00097 (0.00370)	0.00192 (0.00225)	-0.00219 (0.00170)	0.01827** (0.00726)	0.00012 (0.00459)	0.00411 (0.00269)
Background consumption and HARA intercept: $\hat{\omega}$	5.077*** (1.959)	6.761*** (1.992)	8.283*** (2.471)	36.981*** (7.505)	11.838*** (2.932)	15.044*** (3.551)	45.263*** (8.805)

*Notes:* Estimates of parameter differences are inferred from the quasi-maximum likelihood estimates by the delta method. Parameters are separately set for each magnitude among €20, €40, €60, €80 and €160. There are 1,421 observations (203 clusters) for each magnitude. Clustered standard errors in parentheses. Standard errors calculated via the delta method. \*\*\*, \*\* and \* indicate significance at the 1 percent level, 5 percent level, and 10 percent level, respectively.

As we do for the CRRA specification, we use the estimates to predict choices in the 35 questions for both the Present Group and the Delayed Group. Table A 6 presents the marginal effects of allowing parameters to vary with the magnitude: in Row 1-2, we allow  $\delta$  to vary with the magnitude of the decisions (as indicated by the column title), but control other two parameters to be the value estimated from the magnitude of €20; in Row 3-4, we allow  $\alpha$  and  $\omega$  to vary with the magnitude. The results show that the marginal effect of allowing  $\alpha$  and  $\omega$  to vary with the magnitude is comparable with the marginal effect of allowing  $\delta$  to vary. It is consistent with our finding in the Tobit estimation, which implies that the magnitude effect on intertemporal substitutability is at least equally important as the magnitude effect on the discount rate. The fact that the magnitude effect on  $\omega$  is strongly significant is in particular consistent with the story that the degree of asset integration is increasing in the magnitude of rewards.



Table A 6. Marginal Effects of Allowing a Parameter to Vary with Magnitudes in the HARA specification

Magnitude:		€40	€60	€80	€160
Parameter values used in prediction:	$\beta_1, \delta_k, \alpha_1, \omega_1$ (Delayed):	13.9	22.0	25.9	28.0
	$\beta_1, \delta_k, \alpha_1, \omega_1$ (Present):	17.3	22.1	26.0	28.1
	$\beta_1, \delta_1, \alpha_k, \omega_k$ (Delayed):	8.5	19.4	26.1	36.5
	$\beta_1, \delta_1, \alpha_k, \omega_k$ (Present):	10.9	21.7	29.0	40.2

*Notes:* The changes in choices predicted by the HARA nonlinear censored model using the parameter values indicated by the row title compared with  $(\beta_1, \delta_1, \alpha_1, \omega_1)$ , for the two groups separately.  $k$  in the row titles stands for the magnitude in the column title. For instance, the first cell in the first row is the difference between the choices made in the seven decisions with the magnitude of €40 predicted by the model with parameter values  $(\beta_1, \delta_k, \alpha_1, \omega_1)$  and those predicted by the model with parameter values  $(\beta_1, \delta_1, \alpha_1, \omega_1)$ . In other words, it is the marginal effect of allowing  $\delta$  to vary with the magnitude from €20 to €40.

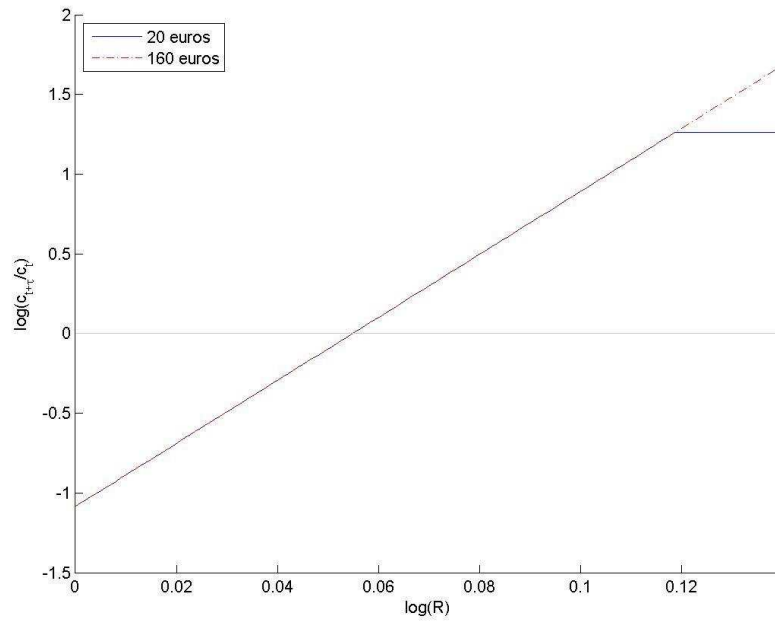
#### Appendix D. A Simulation of the Mental-Accounting Fudenberg-Levine Model

We perform a simulation according to the mental-accounting Fudenberg-Levine model described in Section V. Figure A 1 shows how the dependent variable in our Tobit estimation,  $\ln\left(\frac{z_t + \omega}{z_{t+\tau} + \omega}\right)$ , changes with the independent variable,  $\ln R$ .

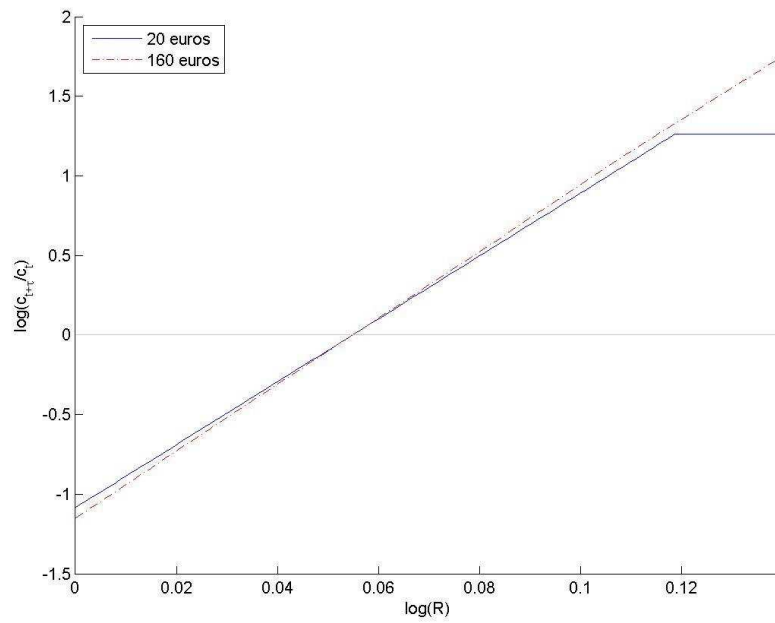
Figure A 1(a) displays the curves in case the true model is specified by equation (1). Note that the slope of the curve (when it is not censored) is  $\frac{1}{1-\alpha}$ . Thus a greater slope stands for a larger intertemporal substitutability. The horizontal-axis intercept is  $-\ln \delta$ . Thus a greater horizontal-axis intercept stands for a smaller discount factor.

Figure A 1(b) displays the curves in case the true model is the mental-accounting Fudenberg-Levine model. When the budget is €20, both rewards are taken as pocket

cash, so the curve is a straight line, as same as predicted by ( 1 ). However, when the budget gets larger, a large fraction of money is saved, and hence the curve is with a greater slope. Therefore, we observe a positive relation between intertemporal substitutability and the size of budget. On the other hand, the horizontal-axis intercept does not change with the stake, suggesting that the model cannot explain the magnitude effect on the discount factor.



(a) Prediction of the model specified by equation ( 1 )



(b) Prediction of the extended Fudenberg-Levine model

Figure A 1. Simulated Relationships between the Dependent Variable and the Independent Variable of the Tobit Estimation

## Appendix E. Decision Forms in Part II

Form 6: Decisions 36 - 42

September 2014	October 2014	November 2014	December 2014
1 2 3 4 5 6 7	1 2 3 4 5	1 2	1 2 3 4 5 6 7
8 9 10 11 12 13 14	6 7 8 9 10 11 12	3 4 5 6 7 8 9	8 9 10 11 12 13 14
15 16 17 18 19 20 21	13 14 15 16 17 18 19	10 11 12 13 14 15 16	15 16 17 18 19 20 21
22 23 24 25 26 27 28	20 21 22 23 24 25 26	17 18 19 20 21 22 23	22 23 24 25 26 27 28
29 30	27 28 29 30 31	24 25 26 27 28 29 30	29 30 31

You are required to allocate **400** tokens among the three dates for each decision:

	Payment A: 22 Sep TODAY	and	Payment B: 20 Oct 4 WEEKS from Today	and	Payment C: 17 Nov 8 WEEKS from Today	22 Sep	20 Oct	17 Nov
36	<u>157</u> tokens at € 0.20 each on 22 Sep	and	<u>43</u> tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€ 31.40	€ 8.60	€ 52.00
37	___ tokens at € 0.19 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
38	___ tokens at € 0.18 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
39	___ tokens at € 0.17 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
40	___ tokens at € 0.16 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
41	___ tokens at € 0.15 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00
42	___ tokens at € 0.14 each on 22 Sep	and	___ tokens at € 0.20 each on 20 Oct	and	<input type="radio"/> 0 <input checked="" type="radio"/> 200 tokens at € 0.26 each on 17 Nov	€	€	€ 52.00

(a) The Present Group

Form 6: Decisions 36 - 42

September 2014	October 2014	November 2014	December 2014
1 2 3 4 5 6 7	1 2 3 4 5	1 2	1 2 3 4 5 6 7
8 9 10 11 12 13 14	6 7 8 9 10 11 12	3 4 5 6 7 8 9	8 9 10 11 12 13 14
15 16 17 18 19 20 21	13 14 15 16 17 18 19	10 11 12 13 14 15 16	15 16 17 18 19 20 21
22 23 24 25 26 27 28	20 21 22 23 24 25 26	17 18 19 20 21 22 23	22 23 24 25 26 27 28
29 30	27 28 29 30 31	24 25 26 27 28 29 30	29 30 31

You are required to allocate 400 tokens among the three dates for each decision:

	Payment A: 22 Sep TODAY	and	Payment B: 20 Oct 4 WEEKS from Today	and	Payment C: 17 Nov 8 WEEKS from Today	22 Sep	20 Oct	17 Nov
36	<input type="radio"/> 200 <input checked="" type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value="0"/> tokens at € 0.20 each on 20 Oct	and	<input type="text" value="0"/> tokens at € 0.20 each on 17 Nov	€ 32.00	€ 0.00	€ 0.00
37	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value="150"/> tokens at € 0.19 each on 20 Oct	and	<input type="text" value="50"/> tokens at € 0.20 each on 17 Nov	€ 16.00	€ 28.50	€ 10.00
38	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value=""/> tokens at € 0.18 each on 20 Oct	and	<input type="text" value=""/> tokens at € 0.20 each on 17 Nov	€ 16.00	€	€
39	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value=""/> tokens at € 0.17 each on 20 Oct	and	<input type="text" value=""/> tokens at € 0.20 each on 17 Nov	€ 16.00	€	€
40	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value=""/> tokens at € 0.16 each on 20 Oct	and	<input type="text" value=""/> tokens at € 0.20 each on 17 Nov	€ 16.00	€	€
41	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value=""/> tokens at € 0.15 each on 20 Oct	and	<input type="text" value=""/> tokens at € 0.20 each on 17 Nov	€ 16.00	€	€
42	<input checked="" type="radio"/> 200 <input type="radio"/> 400 tokens at € 0.08 each on 22 Sep	and	<input type="text" value=""/> tokens at € 0.14 each on 20 Oct	and	<input type="text" value=""/> tokens at € 0.20 each on 17 Nov	€ 16.00	€	€

Submit Decisions

(b) The Delayed Group

Figure A 2. Interface of a Typical Decision Form in Part II