WINPEC Working Paper Series No.E1714
September 2017

Does Time Inconsistency Differ between Gain and Loss? An Intra-Personal Comparison Using a Non-Parametric Designed Experiment

Shotaro Shiba and Kazumi Shimizu

Waseda INstitute of Political EConomy
Waseda University
Tokyo,Japan

Does Time Inconsistency Differ between Gain and Loss? An Intra-Personal Comparison Using a Non-Parametric Designed Experiment

Shotaro Shiba*† Kazumi Shimizu ‡

Abstract

Several studies in the time preference literature have found time inconsistencies (TIs) in both the gain and loss domain. However, their relationship within the same person remains unclear: that is, does an individual who demonstrates TI for gain outcomes also do so for loss? To investigate this relationship, we conducted a non-parametric designed experiment that requires only standard axioms and no parametric specification for people's preferences. In the experiment, we allowed the measurement of TI to depend on character alternatives—such dependency has emerged as a crucial point in recent TI discussions. With these settings, we directly observed TI for gain and loss and found a so-called "future effect" for both outcomes. We also found a positive correlation between the degrees of TI for gain and loss within the same person, irrespective of character alternatives. In addition, in most cases, we found no significant differences between the degrees of TI for gain and loss. These results remained robust even when using another TI measurement. Our findings suggest that people's TI regarding gain and loss may not differ and the source of TI among individuals is common between their preference for gain and loss.

Introduction

It is common knowledge that people occasionally make time-inconsistent decisions; that is, they change their previous decision without situational

*Corresponding author. Graduate School of Economics, Waseda University, 1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, Japan. E-mail: sibasyou2001gmail.com

†Graduate School of Economics, Waseda University, Japan

‡Department of Political Science and Economics, Waseda University, Japan

changes for both gain and loss outcomes. For example, we may incur additional costs to accelerate the delivery of a new television (gain) or knowingly delay the payment of a debt (loss). Such decisions or preferences are the most severe violations in an exponential discounted utility (EDU) model, a standard model used to explore individual's intertemporal choices but requiring the assumption of a constant discount rate, according to which people never change their previous decisions (for a review, see Frederick et al., 2002). Many researchers have empirically and theoretically examined time-inconsistent preferences to improve the understanding of people's time preferences for gains and losses (Ebert & Prelec, 2007; Laibson, 1998; Loewenstein & Prelec, 1992; Sayman & Öncüler, 2009; Theler, 1981).

In this study, we focus on the relationship of time inconsistencies (TIs) with gain and loss. Some scholars have attributed TI to distortion in an individual's perception of the future (Baucells & Heukamp, 2012; Kim & Zauberman, 2009), and if this is true, TI should positively correlate with gain and loss. This is because we probably have a common system to perceive delays in gain and loss. Such correlations enable us to predict an individual's TI using actual but smaller observations. An individual who shows a time-inconsistent preference for gain would naturally tend to do so for loss as well; however, why people behave time-inconsistently remains unclear in the literature. Thus, studying intra-personal relationships with TI will help clarify its source and whether we have a common or distinct TI system for gain and loss outcomes. The findings of this investigation will contribute to the understanding of our "irrationality" in the context of inconsistencies.

TIs for Gain and Loss

TI for gain is a well-studied phenomenon in the literature. In particular, a concept that has gained much attention is the *present effect* (PE), according to which sooner outcomes associated with decreased delays become relatively more attractive than later ones associated with the same change (e.g., Keren & Roelofsma, 1995). For example, we may be indifferent between receiving 100 dollars in a month and 110 dollars in two months, but prefer 100 today to 110 in a month—a mere difference of one month in waiting times. Such behavior gives rise to inconsistencies because it makes people more impatient about future possibilities when time is passing, causing them to change their previous decisions. PEs cannot be supported by EDU given its constant discount rate over time; nevertheless, it can be accommodated in a model that includes the so-called "hyperbolic discounting" or "decreasing impatience," wherein the discount rate decreases with time and people become more patient for far future outcomes (e.g., Loewenstein & Prelec, 1992). More recent studies have proposed the *future effect* (FE), which is the opposite of PE (e.g., Takeuchi, 2011). This effect can be accommodated by an *increasingly impatient* preference in which people become more patient for near future outcomes.

While both effects are contrasting in nature, this does not imply they are incompatible. Since FEs are mostly found when the intertemporal choice is sufficiently short and close in terms of waiting time of alternatives (e.g., one- vs. two-week decisions, not one- vs. two-year decisions), studies have suggested that people are increasingly impatient about sooner intertemporal choices but decreasingly so for later choices. In this sense, PEs and FEs are not constantly exhibited but depend on the conditions of a choice (waiting time of two alternatives). A crucial description of this phenomenon is the inverse S-shaped discounting function (Sayman & Öncüler, 2009; Takeuchi, 2011). Figure **??** illustrates the present values of two alternatives, smaller-sooner (SS) and larger-later (LL), under an inverse S-shaped discounting. The left panel shows PEs in general situations, which is similar to hyperbolic discounting, whereas the right panel depicts FEs when the timings of the alternatives are sufficiently near and close to each other.

Similarly, for loss, Benzion et al. (1989) found a decreasing discount rate, which corresponds to decreasing impatience under the linear utility assumption. Using the same assumption, Thaler (1981), however, found not constant but not decreasing discount rate. This implies that people's impatience for loss does not uniformly decrease. Abdellaoui et al. (2013a) relaxed this assumption and found a decreasing impatience preference for loss outcomes. In addition, they compared the explanation power for various discount functions and identified functions that allowed FE to fit the data better against those that did not. While Abdellaoui et al. emphasized the existence of FEs, no study has directly examined FEs in the loss domain.

Intra-personal Relationship with TI

To the best of our knowledge, Abdellaoui et al. (2013a) is the only study that examines the intra-personal relationship between TIs for gain and loss. Focusing on the difference between time preferences for gain and loss in terms of utility and discount functions, they estimated the parameters of both functions, which were assumed to be sign-dependent, that is, the parameters for loss could differ from those for gain. They measured TI using parameter α of the generalized hyperbolic discount (GHD) function such that $D(t) = (1 + \alpha t)^{-\frac{\beta}{\alpha}}$, where α denotes degree of decreasing impatience or PEs (Loewenstein & Prelec, 1992). Their results showed that decreasing impatience was stronger for loss than for gain at the individual level and their strength had no correlation. However, their measurement α allowed only the existence of PEs, not FEs.

Current Study

Drawing on Abdellaoui et al.'s (2013a) research interest, our study focuses on the difference in individual's TIs. We introduced a measure that allows for FEs and investigated whether (i) FE is also observed for loss and (ii) such TI correlates with that for gain. To deal with these topics more appropriately than previous experimental research, we conducted a laboratory experiment accounting for two essential points, which we explain below.

First, we separately examined TIs for each condition. As mentioned above, TI for gain depends on the condition of choices; that is, FE is mainly found when the intertemporal choice is sufficiently near in the present and the waiting times of the alternatives are adequately close (Figure 1). However, most research, including Abdellaoui et al. (2013a), measured TI by estimating the representative parameter of the discount function, in which the degree of TI is measured not depending on conditions but using parameters across all time periods. Doing so restricts (GHD function) or renders it unclear (e.g., Weibull function). By contrast, in the present design, we separately observed TIs for each condition and provided a TI measurement on each of them. To do so, we offered our subjects for intertemporal choices that were very close to the present (SS arrives within a day), less close to the present (SS arrives in eight days), or far from the present (SS arrives in 29 days), with 1–5 weeks of waiting for LL. We then measured TI for each of these conditions.

Second, our TI measurement did not require a functional assumption. For long, studies have discussed the bias that emerges from assuming a functional form for estimation, particularly in the context of utility (see Frederick et al., 2002). Many studies calculate the discount rate assuming linear utility, which can distort the result (Andersen et al., 2008; Andreoni & Sprenger, 2012a; Takeuchi, 2011). More recent works relax linearity to a concave function, although they are still based on parametric specifications. The specification on utility might be more problematic when comparing gain and loss preferences because the utility functions for gain and loss may differ from each other (loss aversion and sign dependency; see Kahneman & Tversky, 1979). It is generally assumed that utility is concave for gain and convex for loss; however, Abdellaoui et al. (2013a, 2013b) reported linear utility for gain and concave utility for loss. To avoid such a problematic procedure, we employed Rohde's (2010) non-parametric method, in which no functional specifications are required, with minor change in the order of questions.

Focusing on these two methodological points, we found both PE and FE not only for gain but also for loss. In addition, our TI measurement for loss correlated with that for gain and did not significantly differ in most conditions. Importantly, this result was robust even when we used another TI measurement.

Measurement of TI

In this section, we define PE and FE by referencing Prelec (2004) and a degree measurement for them. Let $X = \mathbb{R}$ be a set of outcomes and $T = \mathbb{R}^+$ a set of time periods. We assume that individuals have a time preference $\succsim$ over $X \times T$, which is a weak order. Each element (x, t) of $X \times T$ denotes a delayed outcome, "receiving outcome x at time t." We assume this preference is continuous, monotonic, and impatient. Preferences for gain $\succsim^+$ and loss $\succsim^-$ are subsets of $\succsim$ on $X^+ \times T$ and $X^- \times T$, where X^+ and X^- are sets of gain and loss. For each individual, we assume a reference point of 0, that is, $X^+ = \mathbb{R}^+$. Finally, we define the indifference $\sim$ and the strict relation $\succ$ in a manner commonly done in the literature.

Suppose an individual is indifferent between two delayed outcomes, SS and LL. The standard model for time preference requires that her preference between SS and LL should not be changed with time passage. Formally, for any (x, s), $(y, s+d)$, and s', such that $s > s' > 0$, $(x, s') \sim (y, s'+d)$ must hold if $(x, s) \sim (y, s+d)$ (this is called *stationary*)[1]. However, as explained above, many studies have found that individuals prefer SS (or LL) relatively more when time proceeds. That is, although an individual is indifferent between SS and LL for gain at some time period, she strictly prefers SS (or LL) over LL (or SS) when time $(s - s')$ has passed, because she could not wait for LL any longer (or could wait longer than she had thought in the past period). This shift is reversed for loss because SS becomes more undesirable relative to LL (PE for loss promotes a procrastination).

Definition 1 (Present Effect)**.** For s, s', and d, such that $s > s' > 0$ and

[1]On this issue, Halevy (2015) pointed out that changing the waiting times of alternatives (from s and $s + d$ to s' and $s' + d$ respectively) and the change of the time period of decision (that is, passage from time 0 to $s - s'$) are not the same operations, and thus, non-stationarity and TI are not the same phenomenon. Nevertheless, we consider both as equivalent to allow easier discussion, as most studies do, not attaching much importance to the difference. In other words, we assume time invariance (i.e., the preference does not change over time). However, if this simplification is not accepted, it is noteworthy that all our results remain valid when replacing the words TI and PE with non-stationary and *common difference effect*. See Halevy (2015) for a detailed discussion.

$d > 0$, $\succsim$ exhibits a *PE* for gain from s to s' with delay d when

$$(x, s) \sim^+ (y, s + d) \quad \text{but} \quad (x, s') \succ^+ (y, s' + d),$$

and a *PE* for loss from s to s' with d when

$$(x, s) \sim^- (y, s + d) \quad \text{but} \quad (y, s' + d) \succ^- (x, s').$$

FE can be defined by reversing the relations of these equations.

Prelec (2004) defines *decreasing impatience* for all x, y, and $s < t$ as $(x, s) \sim^+ (y, t)$, which implies $(y, t + \tau) \succ^+ (x, s + \tau)$ for any $\tau > 0$, and $(x, s) \sim^- (y, t)$, which indicates $(x, s + \tau) \succ^- (y, t + \tau)$ for any $\tau > 0$. This is satisfied when $\succsim^+$ and $\succsim^-$ exhibit PE for any SS, LL, and s' under the present assumptions[2]. We note here that our definition of PE is satisfied when and only when $d > d'$ (or $d - d' > 0$) with d' satisfies $(x, s') \sim (y, d')$. This can be easily shown using our assumptions.

Prelec (2004) also compares the decreasing impatience of the two preferences. He defines $\succsim_2$ that exhibits more decreasing impatience than $\succsim_1$ when all accelerating of LL that compensate the effect of decreasing impatience for $\succsim_1$ is insufficient to doing that for $\succsim_2$ (i.e., the individual 2 still prefers SS). In line with this argument, we define a comparison of PE for the two preferences as follows.

Definition 2 (Inter-comparison of TI)**.** Preference $\succsim_2$ exhibits a *stronger PE* for gain than $\succsim_1$ from s to s' with d when
$(x, s) \sim_1^+ (y_1, s + d)$, $(x, s') \sim_1^+ (y_1, s' + d')$, and $(x, s) \sim_2^+ (y_2, s + d)$ hold, but $(x, s') \succ_2^+ (y_2, s' + d')$.
$\succsim_2$ exhibits a *stronger PE* for loss than $\succsim_1$ from s to s' with d when
$(x, s) \sim_1^- (y_1, s + d)$, $(x, s') \sim_1^- (y_1, s' + d')$, and $(x, s) \sim_2^- (y_2, s + d)$ hold but $(y_2, s' + d') \succ_2^- (x, s')$. [3]

An intuitive explanation of this definition is as follows: for fixed time periods s and $s+d$ and common difference $s - s'$, accelerating (or delaying) LL

[2]This definition suggests that adding the same duration of delay for SS and LL makes SS relatively less attractive. The crucial aspect of this definition is that whenever change of delay from t to s offsets the change of outcome from y to x, the change from $t + \tau$ to $s + \tau$ is insufficient to do that for any $\tau > 0$, even though the sizes of change are same in both case. That is, an individual becomes less sensitive to the delay when it is in the far future, which can be interpreted by decreasing the discount rate with time.

[3]This definition is a modified version of the comparison in Prelec (2004). Prelec's definition required the abovementioned relationships even when x of SS for $\succsim_1$ and $\succsim_2$ differ. However, in the present experiment, x of SS does not vary by subject, that is, we focus on the case in which $x_1 = x_2$. Therefore, we use x for simplicity of definition.

$d-d'$ compensates for the *particular* PE of individual 1, although individual 2 still prefers SS over LL in the case of a speed-up (or delay). In other words, accelerating $d-d'$ is enough to compensate for that PE of individual 1, but it does not for individual 2. We note again that this holds if and only if $d'_1 > d'_2$ (or $d-d'_2 > d-d'_1$).

Similarly, we define a comparison between preferences for gain and loss:

Definition 3 (Intra-comparison of TI)**.** $\succsim$ exhibits *stronger PE for loss than for gain* from s to s' with d when
$(x, s) \sim^+ (y^+, s+d)$, $(x, s') \sim (y^+, s'+d')$ and $(-x, s) \sim^- (-y^-, s+d)$ hold, but $(-y^-, s'+d') \succ^- (-x, s')$.

That is, although SS^+ and LL^+ as well as SS^- and LL^- are indifferent for her and the PE of preference for gain is compensated by $d-d'$, it is not sufficient for her preference of loss. For example, suppose an individual strictly prefers ($100, today) to ($120, 1 week) and is indifferent to ($100, 1 month) and ($120, 1 month +1 week). In addition, he is indifferent to (-$100, 1 month) and (-$110, 1 month +1 week), but strictly prefers (-$110, 1 week) to (-$100, today). These preferences imply that he exhibits PE for both gain and loss with such SSs and LLs. Suppose the PE for gain is compensated by an acceleration of three days, that is, ($100, today) $\sim^+$ ($120, 4 days). According to this definition, if he still prefers (-$110, 1 week) with an acceleration of three days to (-$100, today), the PE for loss should be stronger than that for gain.

In this sense, the difference between d and d' can be a measure of a PE degree in our definitions because for d'_+ and d'_- (of $\succsim^+$ and $\succsim^-$ respectively), $d-d'_+ > d-d'_+$ if and only if $\succsim^-$ exhibits stronger PE than $\succsim^+$. Therefore, the value $\delta = d-d'$ represents the degree of TI, and this is comparable (this δ is well-defined under continuity and impatience); i.e., $\delta > 0$ if and only if the preference exhibits PE, and $\delta_i > \delta_j$ if and only if the preference i exhibits stronger PE than the preference j where i and j are 1 and 2 (inter-personal) or $+$ and $-$ (intra-personal). In addition, as mentioned above, many researchers have compared TIs for the two preferences using parameter α of the GHD specification, and under this specification, $\alpha_i > \alpha_j$ holds if and only if $\succsim_i$ exhibits a stronger PE than $\succsim_j$ for all SS, LLs, and $s' < s$ (see Appendix A).

Here, it is noteworthy that the indifferent SS and LL are uniquely determined by a given $(s, d; x)$ under our assumptions, that is, y is unique up to $(s, d; x)$. This is the same for d', that is, d' is also unique up to s', x, and $y(s, d; x)$. Therefore, the measure $\delta = d-d'$ is determined by a given $(s, d, s'; x)$. We write δ as $\delta(s, d, s'; x)$ below. We summarize our measure as follows.

Proposition 1. For any $(s, d, s'; x)$ with $s > s' > 0$ and $d > 0$, the following statements are equivalent:

i) $\succsim$ exhibits PE for $(s, d, s'; x)$

ii) $\delta(s, d, s'; x) > 0$.

In addition, the following are equivalent:

i) $\succsim_i$ exhibits a stronger PE for $(s, d, s'; x)$ than $\succsim_j$

ii) $\delta_i(s, d, s'; x) > \delta_j(s, d, s'; x)$,

where $\succsim_i$ and $\succsim_j$ are $\succsim_1$ and $\succsim_2$ (intra-personal) or $\succsim^+$ and $\succsim^-$ (inter-personal).

Experiment

Participants

We conducted a one-day experiment with 109 students from various departments in Waseda University at 23/01/2017. The subjects were recruited using the university's online portal system and asked to visit our experimental laboratory any time in that day to participate in the study. Of the participants, 69 are male and 40 are female. Most of them are 20 to 24 years old (78 participants), 28 are under 19, and 3 are 25 to 29. Fifteen participants study natural science, whereas 94 do social science or general arts, of whom 7 students study economics. The largest number of subjects study literature (28 participants). The questions included choices of losses and, thus, were all hypothetical. We offered the subjects 800 yen for participation (more than 7 US dollars), which was reasonable since most subjects took about 15–45 minutes to complete the experiment. The instructions and questions were given on the computer screen.

Procedures

Following Rohde (2010), our experimental design comprises two stages. In the first stage, subjects were asked about the amount of money that offsets a delay in receiving 10,000 yen, that is, y^+, such that $(10,000 \text{ yen}, s) \sim^+ (y^+, s+d)$, and y^-, such that $(-10,000 \text{ yen}, s) \sim^- (-y^-, s+d)$. The following is an example of the questions:

Please input a number (X) that would make you feel that B is as good as A.
-A: Receiving **10,000** yen in **92** days.
-B: Receiving __**X**__ yen in **99** days.

We replaced the word "good" and "receiving" with "not good" and "paying" in the loss questions. We repeated these questions with various values of s (92 and 183) and d (7 and 35) .

In the second stage, we asked the subjects about the length of time that offsets changes in the amount of money, that is, d^+ such that (10,000 yen, s') $\sim^+ (y^+, s' + d^+)$ and d^- such that ($-10,000$ yen, s') $\sim^- (-y^-, s' + d^-)$, where y^+ and y^- are the subjects answers in the first stage. The questions were presented using styles similar to those in the first stage.

However, the subjects were not informed that the parameters in this stage (y^+ and y^-) are from their previous answers[4]. Moreover, they were not allowed to go to the previous pages on the computer so that they cannot change these parameters. We set s' to 1, 8, and 29 by controlling for the effect of a weekday.

We asked a total of 32 questions, of which eight were about the amount of money (four each for gain and loss), and 24 were about waiting time (12 each for gain and loss). To control for an order effect, all subjects were divided into two groups and each group was asked a series of questions, progressing from loss to gain related (for one group) and from gain to loss related (for other group) during the two stages ($n = 57$ and 52). Prior to answering the questions, the subjects were asked two training questions.

We provided an example to illustrate our experiment. Suppose in the first stage, a subject stated that 13,000 yen would render "*10,000 yen in 92 days*" and "__ *yen in 99 days*" equivalent and 15,000 yen would make "*10,000 yen in 92 days*" and "__ *yen in 127 days*" equivalent. Then, in the second stage, we asked the subject about the number of days that would allow her to consider "*10,000 yen in 1 days*" equivalent to "*13,000 yen in* __ *days*" and "*15,000 yen in* __ *days*," considering her previous responses of 13,000 and 15,000 yen. As described in the previous section, when the answer to the former question is less than eight days (i.e., $d' < 7$), the subject exhibits PE for $(92, 7, 1)$ because she cannot wait for LL when both the alternatives are available in the present (1 day vs. 8 days) even though she can do that when the alternatives are in the far future (92 days vs. 99 days). Similarly, when the subject's answer is greater than eight, he exhibits FE.

There is extensive literature on the advantages of matching- and choice–based tasks to elicit decision-making behavior (e.g., Bostic et al., 1990). While Andersen et al. (2008) employ choice-based tasks, Takeuchi (2011) and Attema et al. (2010) use matching-based ones; both methods coexist in decision-making theory. Abdellaoui et al. (2013a) adopted choice-based tasks to elicit indifferent time prospects. However, such tasks may not be ap-

[4]To avoid a consistency effect, in the second stage, we randomly ordered the questions under each s' condition.

propriate for the non-parametric method because their grid points are rather rough, and to the best of our understanding, only Bleichrodt et al. (2016) used choice-based tasks in a non-parametric design. Therefore, all elicitations for the questions in the present study are matching based, which also helped maintain subjects' concentration in our experiment. To aid their decision-making and ease the process, we gave each subject access to a calculator and calendar printed with the date and number of days from the day of the experiment.

Analysis

In this subsection, we report the basic results obtained using measure δ. We also introduce additional measures, *hyperbolic factors*, which are an adjusted measure of δ (Rohde, 2010), and α of GHD, a widely used measure. The results are similar across the three TI measures.

TIs for gain and loss

Table **??** summarizes the responses. We eliminated one participant (subject ID 42) owing to unreliable answers[5]. As a result, the total number of observations for δ is 24×108. We included all data for the two ordered groups and corrected certain answers for the following analysis[6]. For the analysis, we adopted a set of general assumptions (i.e., weak order, monotonicity, impatience, and continuity) and used only those observations that satisfied these assumptions. This resulted in 2,096 observations of the 24×108[7]. We discuss the deviation of the non-negligible number of observations from the general assumptions in the Experiment Limitations subsection.

[5]In the first stage, the subject responded with "10,000" for all the questions and in the second, provided confusing numbers.

[6]Although we found significant differences between the two groups in terms of their responses for the loss segment in the second stage, the results do not considerably differ from the main results for each ordered group. In addition, some answers appeared to be incorrectly inputted; thus, we either corrected the answers or excluded them from the analysis. However, the results remain essentially unchanged.

[7]In most cases of violation of these assumptions, participants answered $\leq$10,000 yen, which directly violated impatience (151 answers in the first stage or 453 observations for δ.). In other cases, they answered delay in the second stage, which exceeded the LL date of the first stage (34 observations in addition to those mentioned immediately above). For example, one subject answered $(-10000, 92) \sim^{-} (-12000, 127)$ in the first stage and $(-10000, 1) \sim^{-} (-12000, 180)$ in the second. These two relations violated impatience under transitivity because impatience requires that $(-10000, 92) \succ^{-} (-10000, 1)$ and $(-12000, 180) \succ^{-} (-12000, 127)$, which are inconsistent with the two indifferent relations.

Table **??** describes the ratio of observations that exhibit FEs in each condition. Although most of the subjects demonstrated PEs or time consistency for both gain and loss, there was a sufficient number of observations that showed FE, and these ratios were marginally higher for loss. The table clearly shows certain common tendencies for the gain and loss ratios. First, for both gain and loss, FE was more likely to be observed when the difference between SS and LL was small ($d = 7$) rather than large ($d = 35$). This is consistent with findings in the literature (Sayman & Öncüler, 2009; Takeuchi, 2011). Second, this rate increased when s' changed from eight to one under $d = 7$, although it did not do so under $d = 35$. This is also consistent with the results of previous research. However, it also increased from $s' = 8$ to $s' = 29$, which is inconsistent with existing findings.

Next, we found a positive correlation between the degree of TI for gain and loss. Table **??** presents the calculated δ (and the number of observations) for each condition. Since we focus on the relationship between PEs for gain and loss, we included only those observations that satisfied the gain and loss conditions in the table. Figure **??** depicts the correlations of our measure at the individual level and the lack of clear differences under most conditions. In fact, Spearman's correlation ρ was relatively high in every condition (see the first two rows in Table **??**). Moreover, δ^- did not significantly differ from δ^+ in most conditions (see the results of the sign rank test in Table **??**)[8], although y^+ and y^- in the first stage significantly differed across all conditions (sign rank test; p-values are 0.02, 0.00, 0.00, and 0.00, respectively, in conditions $(s, d) = (92, 7)$, $(183, 7)$, $(92, 35)$, and $(183, 35)$). Thus, we conclude that people's TIs for gain and loss are possibly related.

Hyperbolic Factor

Rohde (2010) provided an adjusted measure for δ used in the present study, which is called a *hyperbolic factor*, and defined it as

$$h(s, d, s'; x) = \frac{1}{sd'(s, d, s; x) - s'd}\delta(s, d, s; x) \quad (= \frac{d - d'}{sd' - s'd}).$$

This measure represents a degree of TI whenever the first term (or its denominator) is positive; Rohde termed this condition *regularity.* He showed that this measure was constant across conditions under the most standard

[8]The sign rank test is less likely to identify a significant difference than the t-test. While we rejected the normality of δ using the Shapiro-Wilk W-test under all conditions (p-values were 0 under every condition), we performed a t-test and obtained almost identical results. We found significant differences for two conditions at the 5% level and in two more conditions at the 10% level.

functional specifications (e.g., exponential discounting, GHD, and quasi-hyperbolic discounting) and coincided with α under GHD specifications (Theorem 8–10 in Rohde (2010)).

Figure **??** presents the results. A total of 693 out of 928 observation pairs satisfied the regularity condition for both gain and loss. According to Rohde (2010), regularity requires that a subject should not be too much decreasingly impatient because the denominator will be negative if and only if d' is too small, which implies a rather large PE.

This figure elucidates that our results for the intra-personal relationship, discussed in the previous subsection, remain valid even when using a hyperbolic factor. This is because they are positively correlated and no significant difference was found in most conditions (see Table **??** and Figure **??**). Indeed, it is noteworthy that the hyperbolic factor varies by condition (Figure 3), indicating that the participants' discounting behavior does not follow the most common discounting functions (Kruskal-Wallis test; p-values for gain and loss are both 0.00). Further, while Abdellaoui et al. (2013a) assumed such functions in most part of their analysis, our measure δ remains valid owing to its non-parametric design.

Parameter Estimation

As mentioned in the previous section, numerous studies have measured TI assuming a GHD function, including Abdellaoui et al. (2013a). Similarly, we adopted α of GHD as a measure and estimated parameters α^+ and α^- of GHD using the non-linear least-square method. We assume the separation of time and outcome in the evaluation function to obtain the following two equations:

$$D(s)u(x) = D(s+d)u(y) \quad \text{and} \quad D(s')u(x) = D(s'+d')u(y)$$

for each condition $(s, d, s'; x)$. Under the general assumptions of discount function D, we derived

$$d' = D^{-1}\Big(\frac{D(s')}{D(s)}D(s+d)\Big) - s'. \tag{1}$$

If we assumed D as the GHD function, d' would be

$$d' = \frac{1}{\alpha}\Big[\frac{1+\alpha s'}{1+\alpha s}\Big(1+\alpha(s+d)\Big) - 1\Big] - s'. \tag{2}$$

Even if we assumed a parametric specification for discounting, the present experiment would be free from specifications on the utility function u because the two intertemporal choices contains the same outcomes, x and y, and thus $u(.)$ is canceled out in equation (**??**). This structure is identical to those in Attema et al.'s (2010) and Takeuchi's (2011) models, so that such experiments are also called "utility-free" methods. We discuss their importance in the Treatment of Utility Function subsection.

According to equation (**??**), parameter β of GHD and, thus, the value of function $D(t)$ cannot be estimated in the present study. Takeuchi (2011) proposed estimating them by deriving the ratio of utility $\frac{u(x)}{u(y)}$ from risk preferences in an expected utility (EU) model. However, the EU specification for risk preference is not reasonable (e.g., Kahneman & Tversky, 1979). In addition, the implicit assumption that the utility functions for risk and time are identical may lead to a bias (Abdellaoui et al., 2013b; Andreoni & Sprenger, 2012a). Since this study focuses on TI, and not the degree of time discounting, we do not follow this method.

The best-fit parameters for α^{+} and α^{-} are 4.30 and 2.57, respectively, for the entire data, implying that time preference for gain deviates more from EDU than that for loss. At the individual level, the medians of α^{+} and α^{-} are $7,79$ (92 subjects) and 5.06 (79 subjects), respectively, considering that certain subjects could not be analyzed because their observations that satisfied our assumptions were too few or the estimator was rather large ($\geq 10^{35}$). These parameters demonstrate positive correlation (Spearman's $\rho = 0.68$, $p = 0.00$) and no significant difference (sign rank test; $p = 0.64$) for subjects whose parameters could be fully estimated (75 subjects). Figure **??** draws this result. The mean values for these subjects are considerably large because some of the subjects reported a high α.

Discussions

In this section, we discuss the advantages of our experiment, which uses a simple measure for the degree of TI and a non-parametric design. The measure and experimental design used in the present study require only a set of standard axioms on preference such that they are free from specification errors. We further detail these aspects and consider the validity of our experiment by referencing the related literature.

Measurement of TI

Our TI measure does not require the assumption of special conditions, such as separability, additivity, Rohde's (2010) regularity, or functional specifications for both the utility and discount functions[9]. By contrast, our assumptions on preferences are only weak order, monotonicity, impatience (widely used axioms in the literature), and continuity, a common technical condition. Such elicitation methods for the discounting function were introduced in the risk preference literature (Abdellaoui, 2000), following which Attema et al. (2010) and Takeuchi (2011) provided new methods for time preference study. Attema et al.'s design incorporates time trade-off sequences and elicits the shape of the discount function using log normalization. The convexity and concavity of their elicited function correspond to PEs and FEs. They measured the degree of participants' TI using manipulation (we explain this in the next paragraph) and reported that 36 out of 55 participants exhibited (or were classified under) FE for near future intertemporal choices. Bleichrodt et al. (2016) used the same measure in the health treatment domain and found that 16 of 63 participants demonstrated the same effect. Similarly, Takeuchi (2011), who proposed another design, found that 362 of 550 observations showed FEs.

In sum, using standard axioms, these studies found strong evidence of FE, which most research using parametric specifications has failed to do. While these studies distinguished PE and FE from revealed choices, we believe there is room for further discussion on their TI measurements. Attema et al. (2010) and Bleichrodt et al. (2016) measured participants' TI by integrating deviations of an elicited function from the linear function (which corresponds to EDU) throughout all conditions; however, this procedure had no theoretical justification and may be disputable given that PE and FE as independent conditions cancel out each in such manipulations[10]. Takeuchi's design offers no inconsistency measurement that is comparable between the inter- and intra-personal level. By contrast, we provided a simple TI measure that allows FEs and can be compared at both the inter- and intra-personal level.

[9]Abdellaoui et al. (2013a) asked their subjects about the present values of multiple outcome prospects, such as "300 in a year and 100 today" or " -300 in a year and 100 today," to estimate parameters and the loss aversion coefficient λ, which we explain in the next subsection. As a result, their analysis assumed the additivity of multiple dated outcomes on evaluations, which is a common but restrictive condition (for a review, see Manzini & Mariotti, 2009).

[10]In fact, for this reason, they also integrated the absolute value of deviations. In this measure, PEs and FEs are identically treated as a deviation from stationarity, such that it represents the size of inconsistency. However, in such cases, PE and FE cannot be identified.

Treatment of Utility Function

In the time preference literature, the curvature of utility function has been a main concern to solve a puzzle that the calculated discount rate (for gain) significantly varies by empirical research assuming linear utility. The rates widely range from negative to more than 300% or 1,000% (see Frederick et al., 2002, for a review). Two important approaches used to control for curvature are estimating parameters using a multiple price list (MPL) of delayed prospects and risky prospects (e.g., Andersen et al., 2008) or a convex time budget (CTB) design (e.g., Andreoni & Sprenger, 2012a). These studies argue that the obtained discount rate is considerably smaller than that with a linear utility assumption. For instance, Andersen et al. (2008) reported a rate of 0.10, although the rate with a linear assumption was 0.25. However, on the one hand, Andersen et al. estimated that parameter θ of the CRRA curvature ($u(x) = \frac{x^{1-\theta}}{1-\theta}$) was 0.74, which is far from linear, and on the other, Andreoni and Sprenger (2012a) showed that the utility function marginally differed from the linear (the estimated curvature of CRRA was 0.08 in the aggregate data and its median was 0.03 at the individual level)[11]. Similarly, Abdellaoui et al. (2013a, 2013b) concluded that a linear utility function for gain at the aggregate level (by their individual data, most participants (74.1%) were classified as linear). In addition, while Andersen et al. (2008) used utility for risky outcome to control its curvature, the latter two research studies suggested that the utility functions for time preference and risk may differ from each other. On the other hand, while Andersen et al. focused only on the single delayed outcome, their studies heavily depend on the additivity of multiple outcomes, which is also debatable (for a review, see Frederick et al., 2002; and Manzini & Mariotti, 2009). In sum, the types of utility function people have for time preference and how studies should treat them in empirical estimations remain unclear, although a specification of the utility function is crucial to estimate a discount function.

Furthermore, assumption of utility function also related to the asymmetry of gain and loss preferences, because the most important phenomenon of such asymmetry, loss aversion and sign dependency, have been generally captured by parameters of specific utility function (Kahneman &Tversky, 1979 [hereinafter KT] for risk and Loewenstein & Prelec, 1992 for time preference). However, in time preference literature, these phenomena are still under-studied so that the validity of KT's specification is not clear. Abdellaoui et al.(2013a) estimated the loss aversion coefficient (λ of KT) at about 1.3, which was almost half of the values presented in Tversky and Kahne-

[11] Andreoni and Sprenger (2012a) described the CRRA function as $\frac{x^{\alpha}}{\alpha}$; thus, we calculated θ using the $1 - \alpha$ of their estimator.

man's (1992) study of risk preference ($\lambda = 2.25$) whereas Tu's (2004) analysis, which adopted a linear utility assumption, was similar to theirs ($\lambda = 2.00$). In addition, Abdellaoui et al. (2013a) reported a concave utility for loss, which was inconsistent with KT's utility under risk. We believe that they are not contradictory because utility under risk and over time are not necessarily the same; however, we should point out that the unity of utilities for risk and time preference has been accepted in certain studies (e.g., Andersen et al., 2008; Takeuchi, 2011) and that the literature lacks reasonable alternatives to KT's function for capturing the two effects. In sum, the topic of utility over time remains underexplored and lacks a stable consensus, especially in the context of loss preference.

Recent studies on non-parametric (or in their terms, utility-free) designs, such as ours, are free from such limitations. In this study, the measurement of δ does not depend on the utility form, and thus, we can avoid specification errors in our analysis. This strong feature is effective even when we assume a functional form for the discount function and measure TI using the α of GHD, which has been adopted by many studies (equation **??**).

Experiment Limitations

Thus far, we have emphasized the validity of our results; however, this study and its results are not free from limitations. First, and most important, is the number of participants who did not conform to impatience and transitivity. These axioms are fundamental in utility theory. Although violation of impatience is not commonly observed in the literature, it is by no means rare. Negative discounting (i.e., preferring a delayed outcome) is mostly found for health, not monetary, outcomes (e.g., Ganiats et al., 2000), but Warner and Pleeter (2001) reported that many older officers might have zero or negative discounting. Prelec and Loewenstein (1998) also found debt aversion, that is, a tendency to pay even before consumption[12]. Some studies have also reported the violation of transitivity (e.g., Roelofsma & Read, 2000).

Second, our finding of a positive correlation with TI was contrary to that of Abdellaoui et al. (2013a). Nevertheless, we believed their result is distorted owing to an error in their measurement α of the GHD function, which did not allow for FEs. On the other hand, our experiment reported a positive correlation between α^+ and α^-. Thus, this contrary result may be

[12]This study does not focus on debt aversion, which is a choice between alternatives with both gain and loss, namely, consuming (gain) with payment (loss). However, according to standard theory or, more specifically, additivity of evaluation functions and impatience, a decision maker would prefer to pay after consumption because post-payment is discounted more than pre-payment.

due to a different reason, possibly the variation in elicitation methods (choice vs. matching) or the specification error in the utility function, as noted above. It could also be attributed to differences in the length of time. While delays of alternative varied from three months to five years in Abdellaoui et al.'s experiment, those in the present study ranged from a day to about 50 days, with the latest alternative being in 218 days. This result does not seem to be in line with that of Abdellaoui et al.: TI for gain and loss may be correlated only when alternatives involve similar delays, and not otherwise (similar to S-inverse discounting). Since some researchers stated that time preference depends on the time horizon (Ebert & Prelec, 2007; Read, 2001), this hypothesis may not be an uncommon one. We leave the exploration of switches in TI from short- to long-term horizons to further research.

The third is the fewer number of observations (15%) that exhibited FEs compared to that in other studies with a non-parametric design (as stated above). A notable difference between existing experimental designs and that in the present study is the order of time-based questions. More specifically, we first offered subjects alternatives that are considerably far from "now," followed by those that were near; by contrast, they adopted a reverse order. We consider our order of questions more natural because this order plays a crucial role in TI (as described in the first section), although this possibly explains subjects' transitivity violation. According to certain psychologists, the longer the delay, the more abstract is the recognition of a delayed outcome (Trope & Liberman, 2010). Possibly, some participants may experience more confusion when they begin with distant rather than near choices. Given that the order of questions affect the validity of non-parametric analysis, investigation of this relationship–a task not yet attempted–is an important research direction.

Conclusions

This study focused on TI for loss and its intra-personal relationship with TI for gain. Thus, we conducted a non-parametric designed experiment requiring minimum preference conditions and introduced a more appropriate measure of TI than those in the literature. Our study make two important contributions. First, we directly observed the FEs of individuals' time preference for loss, which was overlooked by Abdellaoui et al. (2013a). This behavior was most frequently observed when SS and LL were sufficiently close to each other ($d = 7$) and near enough to the present ($s' = 1$), a finding consistent with previous studies on gain outcomes. At the same time, however, they were far from the present ($s' = 29$) frequently, which is inconsistent with existing findings.

Second, we found relatively strong positive correlations between the degrees of TI (δ) for gain and loss. Abdellaoui et al. (2013a) reported no such correlation, although their study assumed functional specifications (and additivity). By contrast, our result was observed in a more plausible setting, and we found not only correlations but also no significant difference between the degrees of TI for most conditions. This result remains robust even when we used other measures for TI (hyperbolic factor h and GHD's parameter α). That is, even if people more heavily discount for delayed gain than delayed loss (Thaler, 1981; Benzion et al., 1989), individuals' tendencies regarding TIs are strongly related between both domains (and may be identical). More specifically, the strength of TI for gain and loss of an individual may be same even when degree of impatience (discount rate) for gain differ for loss.

The results of our study suggest that people experience similar distortions in their delay evaluations, even if the evaluations (discounting) differ from each other. This is possibly consistent with an existing explanation—that TIs are caused by a distortion in people's perception or cognition of both time and delay. It is important that we study the sources of time-inconsistent preferences to understand our rationality and the factors causing a distortion in people's discounting behavior. Thus, this issue has recently become a key research topic in the literature. Many studies have attempted to present the inherent sources of such preferences; these sources might be found in people's minds, which have been evolutionarily hardwired (Baucells & Heukamp, 2012; Dasgupta & Maskin, 2005; Halevy, 2008; Kim & Zauberman, 2009). In conclusion, our results strongly suggest that people may have a common system of evaluation for delayed gain and loss outcomes, which can be associated with their "irrational" behavior. That is, the source of TI may be partly shared or could be identical through the preference for gain or loss.

Appendix A

Here, we show that $\alpha_i > \alpha_j$ if and only if $\succsim_i$ exhibits a stronger PE than $\succsim_j$ for any SS and LL under the GHD specification. This fact was almost proven in Rohde (2010) under the setting $s' > s$.

(Proof). Using the same logic as that in Rohde's (2010, p.133) Theorem 9, we can show that regularity is satisfied under GHD and α must coincide with $h(s, t, s'; x)$ for any $(s, t, s'; x)$ in our setting. From Theorem 4 (Rohde, 2010, p,131), it also follows that the regularity guarantees the equivalence of $h_i(s, t, s'; x) > h_j(s, t, s'; x)$ for all $(s, t, s'; x)$ and the latter statement mentioned above. Therefore, under the GHD specification, $\alpha_i > \alpha_j$ holds if and only if $h_i > h_j$ holds for any condition, and the latter holds if and only if $\succsim_i$ exhibits a stronger PE than $\succsim_j$ for any condition.

Appendix B: An Experiment Sample Material of Instruction (Gain Scenario, Part 1)

Scenario 1: Please imagine the following situation.
Question: You are now able to receive one of the following two alternatives. Please choose one you prefer the most.
-A: "Receiving 1,000 yen in 5 months"
-B: "Receiving _X_ yen in 10 months"
Note: Assume that receiving this money with its delay is guaranteed.
Note: In the actual questions, the values "5 months," "1,000 yen," and "10 months" vary by question.

For example, suppose X is 1,001 or 10,000. In the question above, you were asked whether you could wait for " 1,001 (or 10,000) yen in 10 months" instead of " 1,000 yen in 5 months." You may choose A if X is small or B if it is large.

In this scenario, we are specifically asking you about the amount of money (X) that would make B as good a choice as A.

Example: Select one of two alternatives in the following. Please input in the blank, numbers (X) that would make you feel B is as good a choice as A (you can use the calendar and calculator at your desk).
-A: "Receiving 1,000 yen in 5 months."
-B: "Receiving X yen in 10 months."
Please imagine various X amounts to ensure you consider B as good as A. For example, if you feel that "2,000 yen in 10 months" is as good as "1,000 yen in 5 months," then, input "2,000" in the blank. Similarly, if you prefer "1,000 yen in 5 months" to "1,999 yen in 10 months" or "2,001 yen in 10 months" to "1,000 yen in 5 months," then, please input "2,000."

Acknowledgements

This study is financially supported by the Waseda University Grant for Special Research Projects [Project number: 2016A-001].

References

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science, 46*(11), 1497–1512.

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2013a). Sign-dependence in intertemporal choice. *Journal of Risk and Uncertainty, 47*(3), 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), 2153-?2169.

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

Andreoni, J., & Sprenger, C. (2012a). Estimating time preferences from convex budgets. *American Economic Review, 102*(7), 3333?-3356.

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

Attema, A. E., Bleichrodt, H., Rohde, K. I. M., & Wakker, P. P. (2010). Time-tradeoff sequences for analyzing discounting and time inconsistency. *Management Science, 56*(11), 2015–2030.

Baucells, M., & Heukamp, F. H. (2012). Probability and time trade-off. *Management Science, 58*(4), 831-?842.

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: an experimental study. *Management Science, 35*(3), 270–284.

Bleichrodt, H., Gao, Y., & Rohde, K. I. M. (2016). A measurement of decreasing impatience for health and money. *Journal of Risk and Uncertainty, 52*(3), 213–231.

Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect on the preference reversal of using choice indifferences. *Journal of Economic Behavior & Organization, 13*(2), 193–212.

Dasgupta, P., & Maskin, E. (2005). Uncertainty and hyperbolic discounting. *American Economic Review, 95*(4), 1290–1299.

Ebert, J. E. J., & Prelec, D. (2007). The fragility of time: time-insensitivity and valuation of the near and far future. *Management Science, 53*(9), 1423–1438.

Frederick, S., Loewenstein, G., & O'Donogue, T. (2002). Time discounting and time preference: a critical review. *Journal of Economic Literature, 40*(2), 351–401.

Ganiats, T. G., Carson, R. T., Hamm, R. M., Cantor, S. B., Sumner, W., Spann, S. J., Hagen, M. D., & Miller, C. (2000). Population-based time preferences for future health outcomes. *Medical Decision Making, 20*(3), 263–270.

Halevy, Y. (2008). Strotz meets Allais: diminishing impatience and the certainty effect. *The American Economic Review, 98*(3), 1145–1162.
Halevy, Y. (2015). Time consistency: stationarity and time invariance. *Econometrica, 83*(1), 335–352.
Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263–292.
Keren, G., & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes, 63*(3), 287–297.
Kim, B. K., & Zauberman, G. (2009). Perception of anticipatory time in temporal discounting. *Journal of Neuroscience, Psychology, and Economics, 2*(2), 91–101.
Laibson, D. (1998). Life-cycle consumption and hyperbolic discount functions. *European Economic Review, 42*(3), 861–871.
Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics, 107*(2), 573–597.
Manzini, P., & Mariotti, M. (2009). Choice over time. In P. Anand, P. Pattanaik & C. Puppe (Eds.), *The handbook of rational and social choice* (pp. 239–270). Oxford, UK: Oxford University Press.
Prelec, D. (2004). Decreasing impatience: a criterion for non-stationary time preference and 'hyperbolic' discounting. *The Scandinavian Journal of Economics, 106*(3), 511–532.
Prelec, D., & Loewenstein, G. (1998). The red and the black: mental accounting of savings and debt. *Marketing Science, 17*(1), 4–28.
Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty, 23*(1), 5–32.
Roelofsma, P., & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making, 13*(2), 161–177.
Rohde, K. I. M. (2010). The hyperbolic factor: a measure of time inconsistency. *Journal of Risk and Uncertainty, 41*(2), 125–140.
Sayman, S., & Öncüler, A. (2009). An investigation of time inconsistency. *Management Science, 55*(3), 470–482.
Takeuchi, K. (2011). Non-parametric test of time consistency: present bias and future bias. *Games and Economic Behavior, 71*(2), 456–478.
Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8*(3), 201–207.
Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463.

Tu, Q. (2004). Preference points and loss aversion in intertemporal choice. Working Paper. doi:10.2139/ssrn.644142.

Tversky, A., & Kahneman, D. (1992). Advanced in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

Warner, J. T., & Pleeter, S. (2001). The personal discount rate: evidence from military downsizing programs. *American Economic Review, 91*(1), 33–53.

condition				1st stage			2nd stage		
delay	SS	sign	SS'	Mean	S.D.	Obs.	Mean	S.D.	Obs.
d=7	s=92	gain	s' = 1	10,718.52	866.14	93	4.95	5.96	92
			s' = 8				5.17	6.38	92
			s' = 29				5.08	4.73	91
		loss	s' = 1	10,962.99	1,940.99	77	7.41	10.97	74
			s' = 8				6.47	9.17	72
			s' = 29				8.23	12.04	74
	s=183	gain	s' = 1	10,874.93	1,441.39	92	5.09	6.98	92
			s' = 8				4.80	6.20	92
			s' = 29				5.82	7.28	92
		loss	s' = 1	10,657.43	1,353.68	75	8.09	14.71	74
			s' = 8				8.03	15.45	73
			s' = 29				9.35	15.51	74
d=35	s=92	gain	s' = 1	12,794.59	3,015.71	99	14.99	15.63	99
			s' = 8				13.90	13.97	99
			s' = 29				15.04	13.35	99
		loss	s'=1	12282.34	2449.72	87	17.87	21.84	84
			s' = 8				17.90	21.95	83
			s' = 29				16.12	14.70	82
	s=183	gain	s' = 1	12,841.85	2,983.31	100	16.87	18.09	100
			s' = 8				15.30	16.36	100
			s' = 29				16.21	15.12	100
		loss	s' = 1	12,056.44	2,458.16	87	20.78	31.12	87
			s' = 8				19.97	32.12	86
			s' = 29				19.05	24.10	85

Table 1: Descriptive statistics of participants' answers. The total number of valid subjects is 108. The number under "Obs." denote that of observations satisfying our assumption on preference.

	s = 92,d = 7		s = 183,d = 7		s = 92,d = 35		s = 183,d = 35	
	gain	loss	gain	loss	gain	loss	gain	loss
s' = 1	22%(92)	24%(74)	21%(92)	27%(74)	7%(99)	12%(84)	10%(100)	13%(87)
s' = 8	16%(92)	19%(72)	15%(92)	18%(73)	5%(99)	12%(83)	8%(100)	10%(86)
s' = 29	18%(91)	23%(74)	22%(92)	27%(74)	5%(99)	10%(82)	8%(100)	12%(85)

Table 2: Ratios of observations that exhibited FEs in each condition. The number of observations that satisfied our axioms is presented in parentheses.

d	7						35						
s	92			183			92			183			
s'	1	8	29	1	8	29	1	8	29	1	8	29	total
measure δ													
Spearman's ρ	0.63	0.61	0.52	0.59	0.61	0.61	0.73	0.63	0.47	0.66	0.54	0.42	
p−-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Sign rank	**	n.s.	*	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	
Obs.	73	70	72	72	71	72	82	81	80	86	85	84	928
measure h													
Spearman's ρ	0.54	0.61	0.31	0.4	0.52	0.53	0.73	0.46	0.46	0.62	0.5	0.38	
p−-value	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	
Sign rank	***	n.s.	n.s.	n.s.	n.s.	**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	
Obs.	64	63	36	63	62	45	76	56	21	80	73	54	693
measure α													
Spearman's ρ						0.68							
p−-value						0.00							
Sign rank						n.s.							
Obs.						75							

Table 3: Test statistics of comparison of TIs between gain and loss. In the sign rank test rows, "n.s." indicates "not significant" at the 10%, and ***, **, and * denote difference at the 1%, 5%, and 10% significance level, respectively. The difference in the number of observations is because we only use observations that satisfy our axioms. Irrespective of TI measurement, we found positive correlations between them for gain and loss and no significant difference for most conditions.

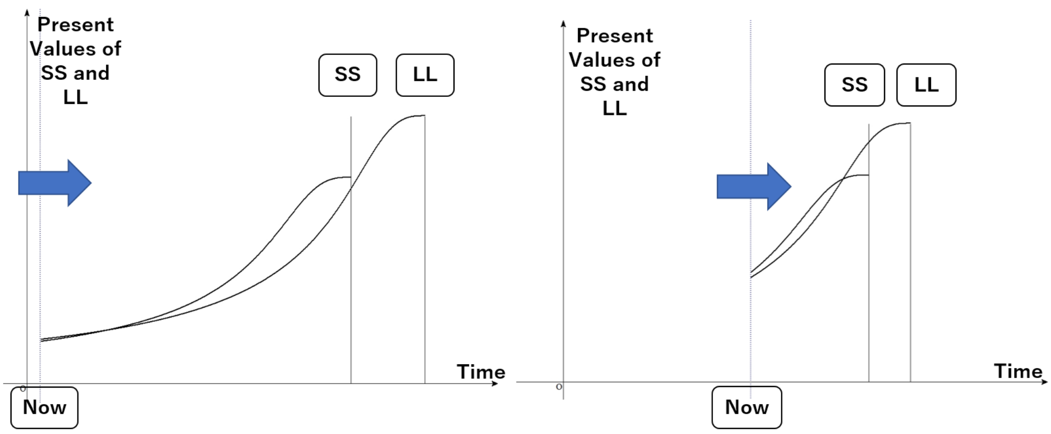

Figure 1: Preference reversals in inverse s-shaped discounting (Sayman & Öncüler (2009, Figure 2, pp.480)). As discount function $D(t)$ has an inverse S-shape in t, its present value with respect to "Now" is described as an S shape in the panels.

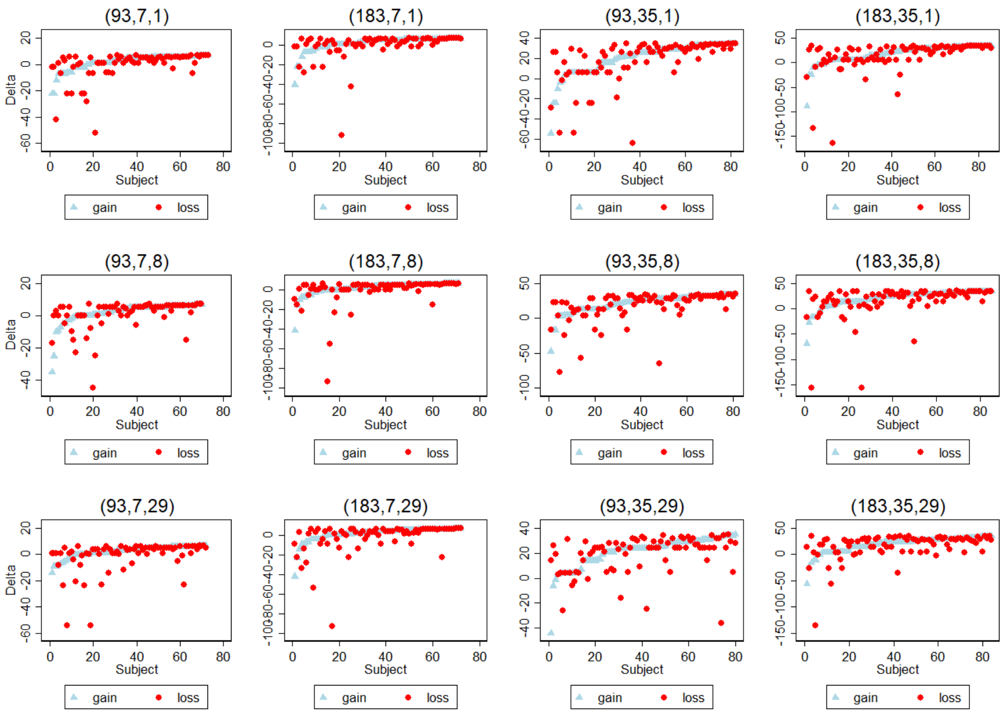

Figure 2: Correlations of δ in each condition. The subjects are ordered by degree of δ^{+}. Each sub-figure corresponds to individual experimental conditions characterized by (s, d, s') at the top of each sub-figure.

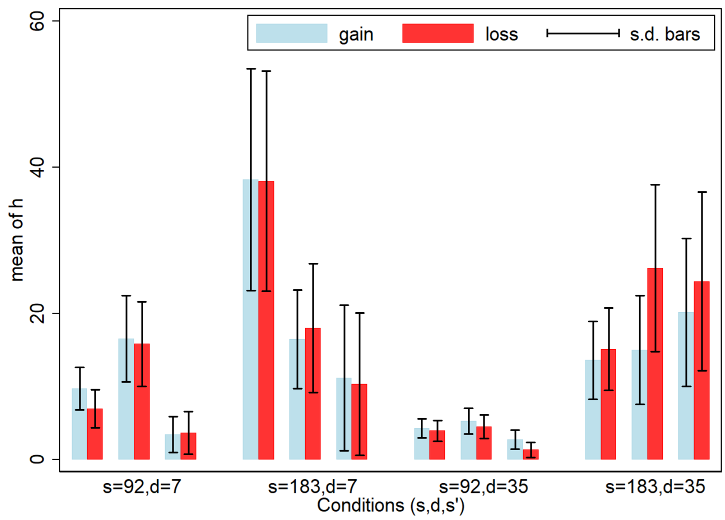

Figure 3: Mean of h with standard error bar of 5%. The bars in each group are put in of $s' = 1$, 8 and 29 from left to right.

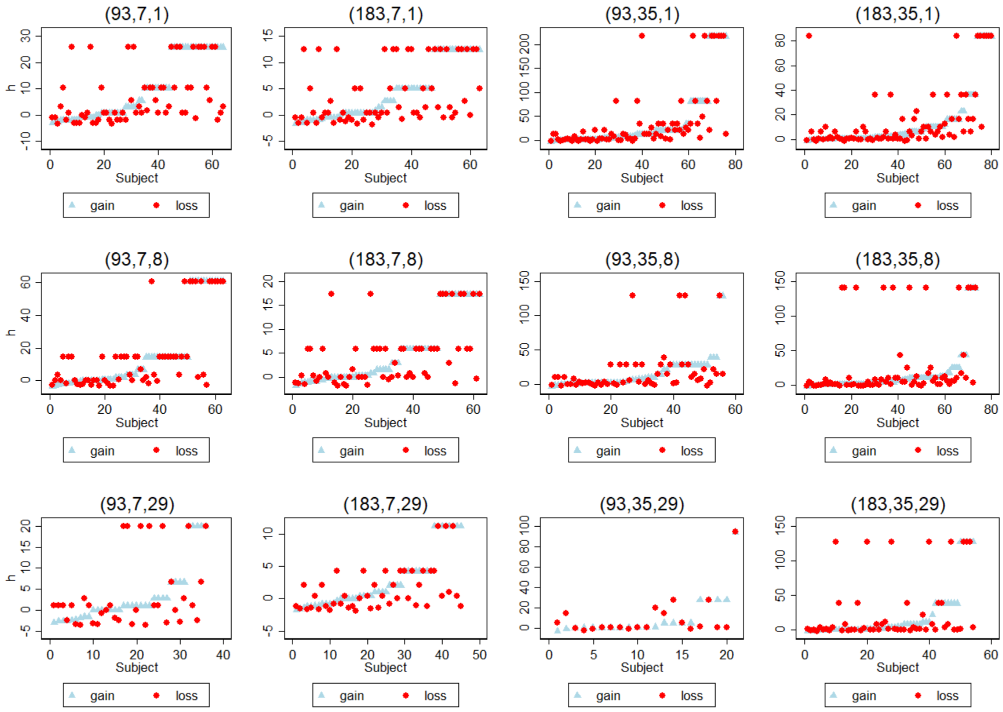

Figure 4: Correlation of h. The subjects are ordered by degree of h^+. Each sub-figure corresponds to individual experimental conditions characterized by (s, d, s') at the top of each sub-figure.

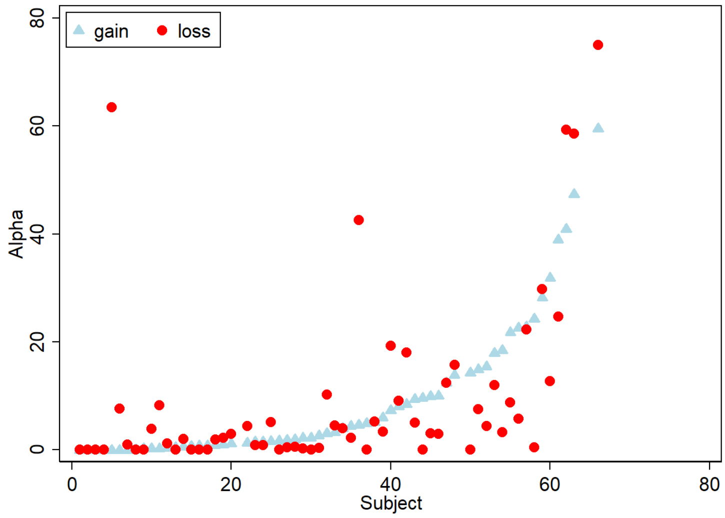

Figure 5: Correlation of α for observations whose parameters are less than 100. The subjects are ordered by degree of α^+.