WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/94732

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

Three Essays on Intertemporal Choice

by

Lisheng He

Thesis

Submitted in partial fulfilment of the requirements for the

Doctor of Philosophy

University of Warwick, Warwick Business School

July, 2017

CONTENTS

List of Tables

List of Figures

Acknowledgements

I would like to thank Professor Daniel Read and Professor Nick Chater for their outstanding supervision and invaluable support. I also thank Dr. Adam Sanborn for his insightful tutorials and warm-hearted support. Their guidance, encouragement and support made it all possible. I am grateful to the China Scholarship Council and the University of Warwick for their financial support. Finally, I am indebted to my families and friends, whose love has brought joy and fun to the lonely and challenging journey.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. This thesis takes a three-paper format, with Chapters 2-4 being independent but conceptually connected working papers.

I declare that the work presented in the thesis (including literature reviewed, experimental design, data collection, data analysis, and writing up) was carried out by myself under normal supervision by Professor Daniel Read and Professor Nick Chater, except in the case below:

- Chapter 1 was partly published in a book chapter for which I served as the third author. See the details of this publication below.
- Chapter 2 was written in collaboration with Marc Scholten (Universidade Europeia, Portugal), Kenneth Lim (WBS), Adam Sanborn (Warwick Psychology) and Daniel Read (WBS).
- Chapter 3 was written in collaboration with Daniel Read (WBS) and Nick Chater (WBS).
- Chapter 4 was written in collaboration with Daniel Read (WBS).

Publication list

Read, D., McDonald, R., & **He, L**. (in press). Intertemporal choice: Choosing for the future. In A. Lewis (Ed), *The Cambridge Handbook of Psychology and Economic Behaviour* (2nd ed.).

Abstract

This thesis focuses on the cognitive processes of intertemporal choice. Chapter 1 is an introductory chapter, laying out the economics standard of intertemporal choice, the environmental complications and cognitive factors that drive the departures from rational intertemporal choice and finally the approach taken in the thesis.

Chapters 2-4 are three empirical studies. Chapter 2 focuses on the evaluation rule of intertemporal choice. Three different evaluation rules have been proposed: alternative-based, attribute-based and hybrid rules. We contrast different evaluation rules by running a comprehensive model comparison in intertemporal choice by involving fifteen candidate models (eight alternative-based, one hybrid and six attribute-based), three stochastic specifications, and 225 data sets taken from the existing literature. Results lend strong support to the class of attribute-based models, especially the family of the tradeoff model, for intertemporal choice.

Chapter 3 studies the attention effects on intertemporal choice. Behavioural theories and experimental studies usually assume an option-wise attention effect on value-based decision making: When an option is focused attention on, the option is given additional weight in the making of decision. Beyond the option-wise attention effect, the study in Chapter 3 reveals a component-wise attentional effect: When each component (or the single value of an attribute in an option) receives attention, it is given additional weight independently. Further comparisons between experiments suggest a probable co-existence of the component-wise and the attribute-wise attention effects, the latter of which is that the comparison along an attribute receives additional weight when focused attention on, on intertemporal choice. The study also demonstrates robust background contrast effects on intertemporal choice.

Chapter 4 focuses on a controversial topic: the pattern of impatience concerning the near vs. the far future (i.e., decreasing impatience, increasing impatience and constant impatience). The study tests two ways to look through the conflicting results in the literature. The first is a design bias when pairs of intertemporal choice items are used to detect the aggregate pattern of impatience. This method makes an implicit assumption that the undetected patterns are homogeneous to the detected and thus generalises the detected patterns to the undetected ones. The present study is the first to test the homogeneity assumption and the results suggest a design bias. The second is an order effect on the detected pattern of impatience, relating to the background contrast effect. Taken together, the two findings could reconcile much variation in the detected pattern of impatience in the literature.

Chapter 5 is a general discussion. I discuss the results from the preceding chapters and implications on theory development. Further discussion on extensions to other domains of intertemporal choice concludes the thesis.

CHAPTER 1 INTRODUCTION

Many of our choices have consequences for the future. For example, we choose to enter a university for a prosperous career in the future. We save now to buy a house (or anything else) in the future. We buy a car now with a monthly instalment in the future. We exercise for future health, etc. For a society, decision making often has more temporally distant consequences for future generations, such as the Paris Agreement on climate change 2015. In such cases, a crucial point is how future benefits are evaluated in relation to immediate costs and how people make tradeoff between consequences in the near future and consequences in the far future.

When a decision involves such an intertemporal tradeoff, it is called an intertemporal choice. For decades, intertemporal choice has been intensively investigated in psychology, economics and management science (for a historical overview, see Loewenstein, 1992). In the abundant literature, there are several different lines of research in this filed. For example, some studies investigated how the degree of impatience in intertemporal choice is related to individual differences in cognitive and personality traits (e.g., Dohmen, Falk, Huffman, & Sunde, 2010; Enzler, Diekmann, & Meyer, 2014; Reimers, Maylor, Stewart, & Chater, 2009; Shamosh et al., 2008).[1] Some compared the degree of impatience in intertemporal choice of different goods, such as money, health, food, drinks and working/leisure hours (e.g., Augenblick, Niederle, & Sprenger, 2012; Chapman, 1996a; 1996b; Ebert, 2010; Estle, Green, Myerson, & Holt, 2007). Some attempted to develop better ways to elicit intertemporal preference (e.g., Andersen, Harrison, Lau, & Rutström, 2008; Attema, Bleichrodt, Rohde, & Wakker, 2010; Coller & Williams, 1999; Toubia, Johnson, Evgeniou, & Delquié, 2012). Some others attempted to find out the models that offer better descriptive accuracy to intertemporal choice (e.g., Cavagnaro, Aranovich, McClure, Pitt, & Myung, 2016; Dai & Busemeyer, 2014; Scholten & Read 2006; 2010; Scholten, Read, & Sanborn, 2014; 2016). Some studies suggested that intertemporal choice, as well as many other types of judgment and decision making, is malleable to a variety of normatively irrelative factors (e.g., Magen, Dweck, & Gross, 2008; Lerner,

[1] Following Fisher (1930), I regard "impatience" as a synonym of time preference, which can be either rational or irrational although in the psychology literature, impatience is always used as an indicator of irrational reluctance to wait for a larger but later reward.

Li, & Weber, 2013; Loewenstein & Prelec, 1992; Read, Airoldi, & Loewe, 2005; Read, Frederick, Orsel, & Rahman, 2005; Read, Olivola, & Hardisty, 2016; Scholten & Read, 2013; Wu & He, 2012).

Despite the increasing popularity, there is a lack of an agreed normative basis for the empirical analysis of intertemporal choice (see Coller & Williams, 1999). Particularly, several key claims are unclear. First, Samuelson's (1937) discounted utility (DU) model has been repeatedly mentioned as the normative model for intertemporal choice while Samuelson himself explicitly stated that his model was not considered as the normative model and did not provide any axiomatic analysis. Second, when comparing the discount rates (or, more generally, the degree of impatience) for different commodities (usually between money and non-monetary outcomes), many researchers hold the null hypothesis that there should be a single discount rate that governs all commodities per the DU model (e.g., Chapman, 1996b), while the DU model itself does not make such an assumption. Third, many researchers claimed that participants in their experiments exhibited excessive discounting in the intertemporal choice of monetary outcomes, compared with the interest rate available in the market (see Frederick, Loewenstein, & O'Donoghue, 2002), while the normative rationale for an association between people's intertemporal preference of consumption and the rate of interest available in the market is rarely established.

To make sense of these conceptual claims, this introductory chapter firstly draws attention to the forgotten economic basis for rational intertemporal choice by laying out the normative accounts for intertemporal preference for two different circumstances: (a) optimal intertemporal allocation of consumption and (b) optimal intertemporal choice. *Intertemporal allocation of consumption* is concerned with how people allocate a fixed bundle of resources to different time periods for consumption so as to maximise their overall utility. By contrast, *intertemporal choice* is concerned with choices between two or more options (such as investment opportunities), which produce different bundles of resources (such as streams of incomes). [2] With the normative accounts, I shall revisit the main findings from intertemporal choice research and discuss a list of environmental factors and cognitive factors that could

[2] Note that this intertemporal choice allows individuals to borrow from and save in a market. The borrowing and saving opportunities are sometimes called *intertemporal arbitrage* in the literature, which is regarded as confound of the time preference for consumption (see Coller & Williams, 1999; Frederick et al., 2002). However, they are the key concepts in Fisher's (1930) normative framework.

lead to empirical departures to the normative predictions. At the end of this chapter, I also briefly explain the approach taken in the thesis.

1.1 Normative Models

The formal analysis of intertemporal choice dates to Irving Fisher (1910; 1930). With the insights from his precedents, Fisher identified six personal characteristics that could shape one's degree of impatience in intertemporal choice (or equivalently intertemporal preference). They are (1) foresight, (2) self-control, (3) habit, (4) expectation of life, (5) concern for the lives of other persons, and (6) fashion. These factors are still among the core interests in the study of intertemporal choice from both economic and psychological perspectives (Frederick et al. 2002; Read, 2004; Read, McDonald, & He, 2016).

Fisher's analysis went far beyond a mere list of these characteristics. Crucially, he assumed that preference of the intertemporal allocation of resources should be influenced by the individual's current consumption circumstance and their expectation of future consumption circumstances. For example, a university student often does not have stable income at present but expects to be better off in the future. She will probably give more weight to the present consumption than future consumption. Thus, the student will show a high degree of impatience. When the same person is well-off with a decent salary at present but expects to get retired in the future without stable income, she will probably give more weight to her future consumption, showing a low degree of impatience. More starkly, the decision maker may even weigh future consumptions more than the present consumption if she is very well-off by now but expect bad financial circumstances in the future.

Based on the crucial assumption on the dependence of actual time preference on background consumption, Fisher (1930) offered a formal framework to analyse rational behaviour in such intertemporal situations where people can reallocate their consumption by borrowing and lending in a market without transaction costs, which gives rise to the net present value (NPV) model as the normative model for intertemporal choice (between tradable goods, especially money).

1.1.1 Converging interest rates

Fisher's (1930) analysis started with how interest rates in a capital market converged. I present a brief illustration of his formal analysis of the converging interest rates with the simplest setup here. Suppose there are only two players, A and B, and that the income streams and consumption streams only last for two periods, t_0 and t_1.

Each of the players is endowed with fixed incomes at the two periods: Player A is endowed with x_0 at period t_0 and x_1 at period t_1. Player B is endowed with y_0 at period t_0 and y_1 at period t_1. Suppose that the incomes are perishable and must be consumed at the same time as they are earned.[3] If lending and borrowing are not possible, their consumption streams, denoted by $\{C_{A0}, C_{A1}\}$ for A and $\{C_{B0}, C_{B1}\}$ for B respectively, should be identical to their income streams (i.e., $C_{A0} = x_0$, $C_{A1} = x_1$, $C_{B0} = y_0$, $C_{B1} = y_1$).

If the two players can borrow from or lend to each other without transaction costs, do they want to reschedule their consumption by borrowing or lending? To answer this question, a key concept is the marginal rate of intertemporal substitution (Frank, 2008, pp.158). The marginal rate of intertemporal substitution (MRIS) is the number of units of consumption in the future (e.g., t_1) one would be *just* willing to exchange for 1 unit of consumption at present (i.e., t_0). Mathematically, it is the absolute value of the slope of the intertemporal indifference curve at a given point. As shown in Figure 1.1, at the endowment, player A's MRIS is R_A, which means that player A is willing to borrow 1 unit for consumption at t_0, at the cost of R units of consumption at t_1 as long as $R < R_A$. Similarly, player B's MRIS at the initial endowment is R_B, which means that B is willing to lend 1 unit at t_0 as long as she can get a return of R units at t_1 if $R > R_B$. Without the loss of generality, suppose $R_A > R_B$. Then player A and player B can reach an agreement on the borrowing-lending scheme with the intertemporal substitution rate R, as long as R satisfies $R_B < R < R_A$. That is, player A borrows 1 unit from player B at t_0 and pay back R units at t_1. After this lending-borrowing scheme is arranged to be implemented, player A's expected consumption at t_0 increases to $(x_0 + 1)$ and his expected consumption at t_1 decreases to $(x_1 - R)$. By contrast, player B's expected consumption at t_0 decreases to $(y_0 - 1)$ and her expected consumption at t_1 increases to $(y_1 + R)$.

[3] Note that the incomes in Fisher's (1930) terminology is not necessarily monetary outcomes. They are usually directly consumable goods.

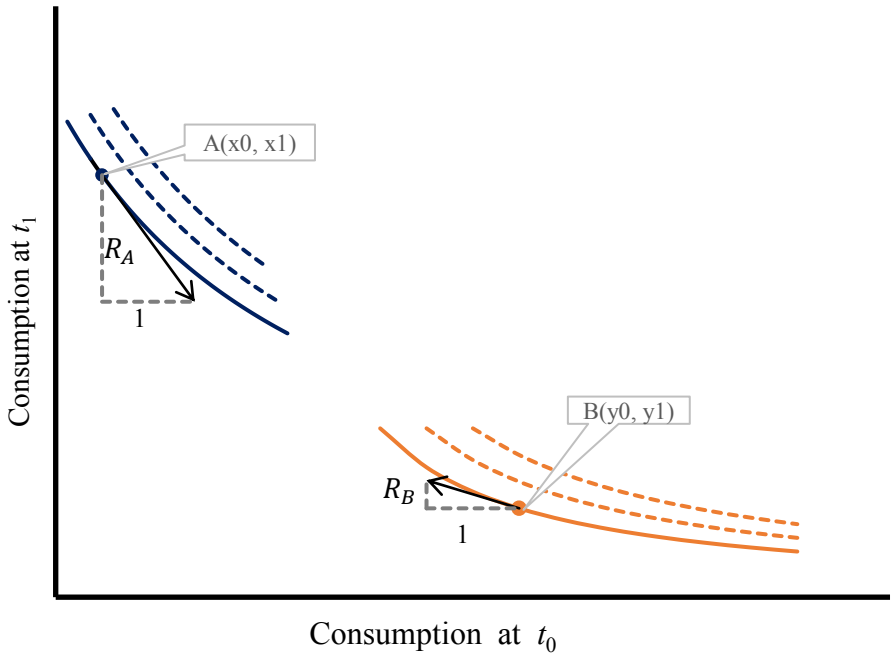

Figure 1.1. Indifference curves and marginal rates of intertemporal substitution (R_A and R_B) of the two players at their initial endowments. Both players are willing to trade with each other at certain rate (to reach a better indifference curve) until their marginal rates of intertemporal substitution converge.

The changes to their expected streams of consumption should influence their MRIS. Specifically, player A's MRIS will decrease and player B's MRIS will increase. Thus, the difference between the two players' MRIS becomes smaller. This lending-borrowing scheme will iteratively continue, as long as player A's MRIS is still larger than player B's, until their marginal rates of intertemporal substitution converge ($R_A^* = R_B^*$), reaching a stable market equilibrium. The converging MRIS becomes the intertemporal substitution rate, R^*, in the market equilibrium ($R^*_{0\to1} = R_A^* = R_B^*$). In other words, the resulting interest rate in the market is indeed jointly determined by the time preferences of the players in the market.

The illustration above can be generalised to situations where there are any number of players in the market (Fisher, 1930). Importantly, when the number of players is large enough, each player becomes negligible in the market, which means that an individual player's time preference of consumption will have a negligible effect on the interest rate at the market equilibrium. This lays the basis for the analysis of individuals' optimal intertemporal allocation of consumption and rational

intertemporal choice between different bundles of resources, which are the focus of Sections 1.1.2 and 1.1.3 respectively.

1.1.2 Intertemporal allocation of consumption

In a market where there are a very large number of players, each player will have access to a stable per-period interest rate of $r_{0\to1}$ (where $r_{0\to1}$ = $R^*_{0\to1}$ -1) between periods t_0 and t_1, where $R^*_{0\to1}$ is the marginal rate of intertemporal substitution (MRIS) between t_0 and t_1 in the market equilibrium. Thus, with a bundle of resources, one could allocate any amount to each point of time along the dashed-straight budget line as shown in Figure 1.2. Intuitively, individuals' allocation depends on how differently they value the consumptions at the two periods. Suppose someone only care about her consumption at t_0. She will allocate all the resources (i.e., X) to consumption at t_0. By contrast, if someone only care about the consumption at t_1, she will allocate all the resources (i.e., $X(1+r_{0\to1})$ or $X\ R^*_{0\to1}$) to consumption at t_1. However, most people are not so extreme and usually prefer to spread their consumption over time, which is also the crucial assumption in Fisher's (1930) framework.[4] In other words, when the consumption concentrates on only one period, they are willing to sacrifice a large sum of consumption from that period for a small sum of consumption at the other period. Thus, their indifference curve will be convex as shown in Figure 1.2. Correspondingly, the best indifference curve that they can attain (representing the maximum utility from consumption) is the one that the budget constraint line is tangent to. In other words, people should make their allocation decisions according to the only point of contact (POC) between the curve and the budget constraint line. Any indifference curve that is above (or better than) this indifference curve is unattainable with this budget constraint line.

[4] Obviously, this is also rational in terms of the fitness for survival.

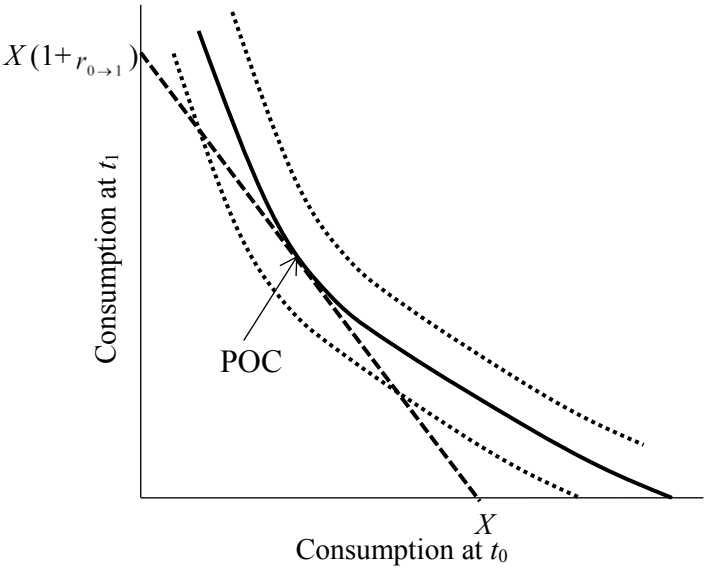

Figure 1.2. The optimal allocation of consumption to the two periods.

Because of the convexity of the indifference curve, the marginal rate of intertemporal substitution (MRIS) between the two periods depends on the allocation of consumptions of the resources to the two periods. Accordingly, the actual *observed* intertemporal preference or MRIS is variable, contingent on the pattern of background consumptions the individual is endowed with. For example, if someone has a large sum of consumption at t_0, but a small one at t_1, she is probably willing to sacrifice a large sum at t_0 in exchange for a much smaller sum at t_1. Fisher (1930) regarded pure time preference as the MRIS of the indifference curve when the allocations to the two periods are equal. Defined in this way, Fisher's (1930) *pure* time preference for consumption is often not observable and is different from the observed time preference researchers' observation in the field or experiments.

1.1.3 Intertemporal choice

Researchers are particularly interested in intertemporal choice when two or more options are offered. According to Fisher (1930), in a perfectly competitive capital market, options can be evaluated and compared according to the net present value (NPV). In the NPV model, a stream of incomes can be evaluated by being translated into an equivalent value at present.[5] For example, the NPV of the stream of

[5] Although the *present* time point is routinely used as the reference point, it does not matter if a future time point is used to calculate the equivalent value.

x_0 at t_0 (the present) and x_1 at t_1 (a future point) can be represented by $NPV = x_0 + \frac{x_1}{1 + r_{0\to1}}$, where $r_{0\to1}$ is the per-period interest rate between t_0 and t_1.

It is not difficult to derive that people should choose the option that maximise the net present value in a market with costless, stable and accessible borrowing and lending opportunities, regardless of individuals' time preference for consumption, because the option with the highest net present value offer a dominant budget constraint line. Take the choice between option X and Y in Figure 1.3 for example. Although Option X is preferred to Option Y at their initial endowments (according to the orange indifference curve), the budget constraint offered by Option X is dominated by that offered by Option Y (according to the dashed budget constraint lines). So, with the given interest rate in the market, decision makers should choose Y instead of X, because Y offers a better net present value or budget constraint than A does and thus attains better indifference curves (see the blue indifference curve).

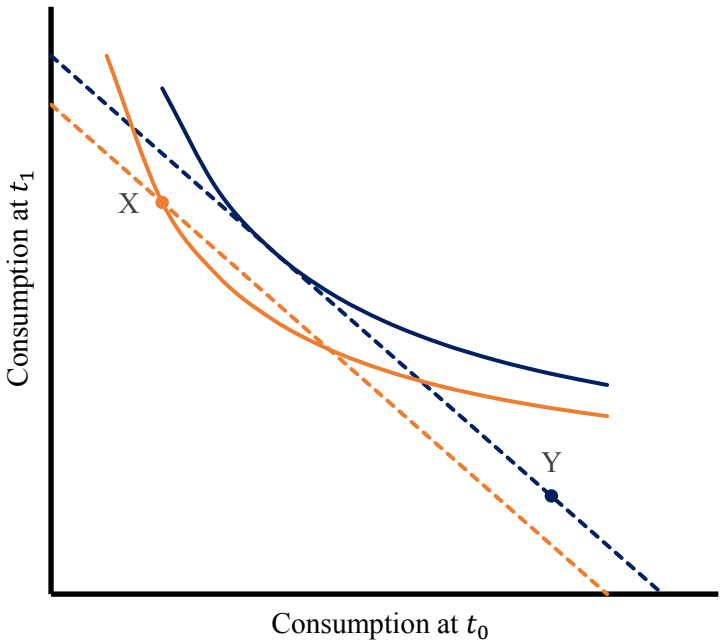

Figure 1.3. The budget constraint lines of two options with different net present values. Option X is preferred to Option Y at their initial endowments, but the budget constraint line from Option Y (blue dashed) dominates the counterpart from Option X (orange dashed) according to the interest rate in the market.

The NPV model can generalise to multiple periods with the inter-period intervals of the same length. Individuals' optimal choice is the option that maximise the NPV according to the interest rates available in the market:

$$NPV = x_0 + \sum_{t=1}^{T} x_t d_{0\to t}(t) = x_0 + \sum_{t=1}^{T}\left(x_t \prod_{\tau=1}^{t} \frac{1}{1+r_{\tau-1\to\tau}} \right),$$

where x_t is the amount of resources (often money) available at time *t*, $d_{0\to t}(\cdot)$ is the discount function for the interval between the present (i.e., a delay of 0) and time *t*, $r_{\tau-1\to\tau}$ ($\tau \geq 1$) is the per-period interest rate over the interval between two consecutive periods τ-1 and τ. Note this analysis uses discrete time rather than continuous time.

1.1.4 Forms of discount functions

In this section, I discuss the forms of discount functions in both the NPV model and the DU model over multiple periods.

The net present value (NPV) model. Fisher (1930) did not provide a general form of the discount function. However, research and practice in economics and finance often assumes constant interest rate over time, i.e., $r_{\tau-1\to\tau} = r$ for all $\tau \geq 1$ (e.g., Brealey, Myers, & Allen, 2012; Frank, 2008, p.156; Hey, 2003), although constant interest rate does not have a strict normative basis (Fisher, 1930). Thus, the net present value (NPV) model becomes,

$$NPV = \sum_{t=0}^{T} x_t \left(\frac{1}{1+r} \right)^t,$$

where *r* is the constant per-period interest rate available in the market. The corresponding per-period discount factor is $\delta = \frac{1}{1+r}$ and the per-period discount rate is 1 - δ.[6]

The discounted utility (DU) model. In terms of the intertemporal allocation of consumption, Paul Samuelson's (1937) discounted utility (DU) model has been long regarded as the normative model:

[6] Keynes (1936) discussed commodity-specific interest rates when the exchange rate between commodities changes over time. However, this is out of the scope of the normative account. In the normative account, a strong assumption that the exchange between commodities keeps constant over time is made.

$$U = \sum_{t=0}^{T} u(c_t)\left(\frac{1}{1+r_c}\right)^t,$$

where r_c is the constant per-period personal interest rate for consumption. The corresponding per-period discount factor is $\delta_c = \frac{1}{1+r_c}$ and per-period discount rate is $1 - \delta_c$.

Strotz (1955) shows that a constant interest rate for consumption in the DU model is necessary to achieve time-consistent allocation of consumption over time.[7] Otherwise, people would keep changing the allocation of consumption over the passage of time. For example, as shown in Figure 1.4, a decreasingly impatient person gives special weights to temporally proximal selves and over-consume the resources and thus leave less for future selves. When future selves come closer in time, one of the "future" selves become the "present" one, she is going to re-evaluate and again give special weights to the present and temporally close selves and over-consume. When this iterative process happens for multiple selves, the far-future selves will get almost nothing. Koopmans (1960) provides a formal axiomatization for the constant interest rate in the discounted utility model.

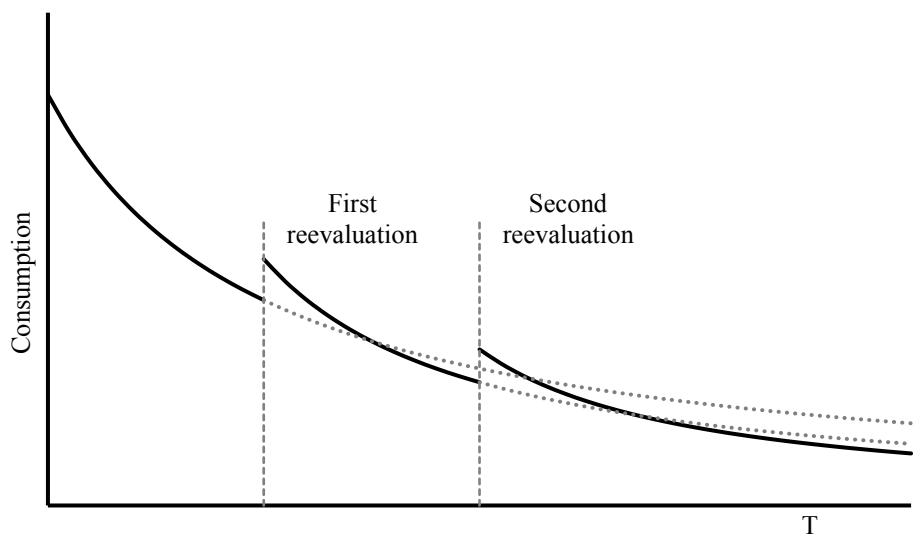

Figure 1.4. Time-inconsistent allocations of consumption over time (Strotz, 1955).

[7] Note that this is not incongruent with Fisher's (1930) argument that the degree of impatience should be influenced by background consumption, because Fisher's (1930) observed time preference is defined on the objective amounts of consumption but the time preference (or discount rate) in Samuelson's (1937) DU model is defined on the subjective utility from consumption.

In terms of the interest rate for different goods, Samuelson (1937) does not discuss whether there is a single interest rate r_c governing all different types of goods or there are good-specific interest rates. However, I argue that a unity of the interest rate for the consumption of different goods should hold as a postulate for the DU model. Otherwise, a cross-modal intertemporal choice could result in dynamic inconsistency in intertemporal choice (see Read & van Leeuwen, 1998; Read, Loewenstein, & Kalyanaraman, 1999).[8] For example, suppose an apple is worth 10 utils, a chocolate is worth 15 utils and the discount rate for the utility of consuming a chocolate is larger than the discount rate for the utility from consuming an apple. Considering a choice of a chocolate and an apple available at the time point T20 (see Figure 1.5). At T0, the apple is preferred to the chocolate but, the preference is reversed when the time of decision comes closer to the time of consumption. For example, at T15, the chocolate is preferred to the apple.

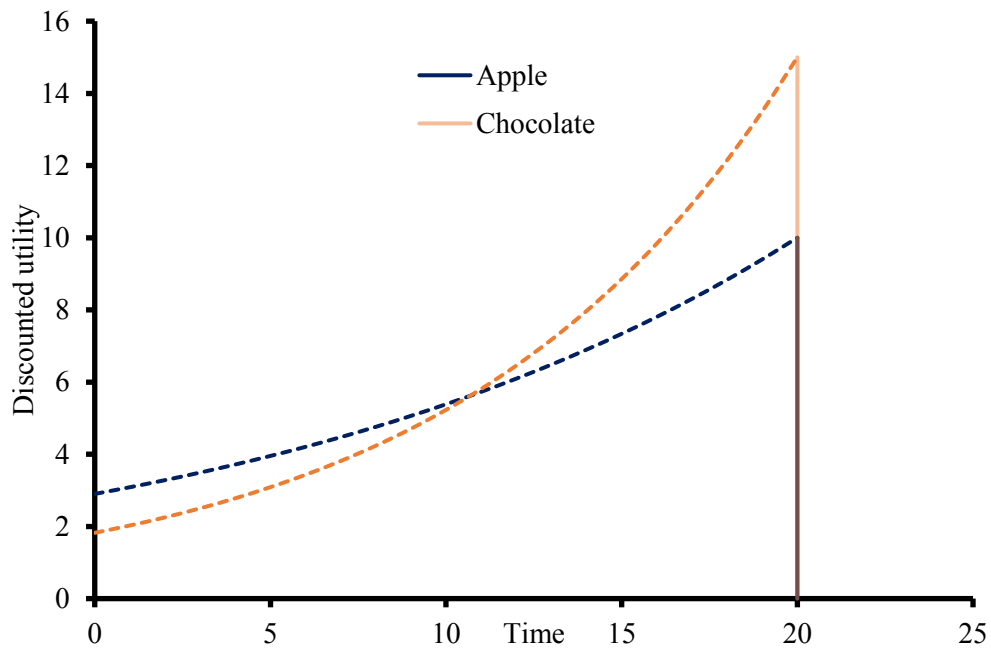

Figure 1.5. Time-inconsistent choice between different goods when discount rates differ across goods.

[8] Another way of interpreting the time-inconsistent preference of the apple and the chocolate is that the consumption of the chocolate is for immediate enjoyment but the consumption of the apple is for future health. However, this alternative explanation does defect the illustrative power of different discount rates for the consumption of different goods.

Differences between the NPV and the DU models. There are two salient distinctions between the NPV model and the DU model. First, The NPV model involves the discount of the *market value* of bundles of resources, as long as they can be traded in the market, but the DU model involves the discount of the *utility* from consumption of resources. Second, the interest rate in the NPV model is the rate available in the market, while the interest rate in the DU model is a personal interest rate of utility from consumption and is unrelated to the interest rate in the market. Put in another way, the interest rate in the NPV model is universal to all participants in the same market but the interest rate in the DU model could be person-specific. However, many studies on intertemporal choice failed to make a distinction between them and thus making mistaken claims.

The use of the NPV or DU model as the normative model depends on the objective of a study. If a study is to investigate whether people are excessive discounters compared with the interest rate available in the market, the NPV model should be used as the normative standard. However, it is important to note that when using NPV, we assume that the incomes or goods are tradable in a market (e.g., money, food). If a study focuses on the consumption of incomes and goods, which is always the case in the literature, DU should be used as the normative standard. Researchers could compare people's discount rates from the DU model. However, we should be cautious about the assumptions we make for the utility function. For convenience and simplicity, the utility of consumption is always deemed equivalent to the raw amount of consumption (e.g., Charlton & Fantino, 2008; Estle et al., 2007; Odum, Baumann, & Rimington, 2006; Odum & Rainaud, 2003; Tsukayama & Duckworth, 2010; Ubfal, 2016). This is problematic for many cases due to reasons such as satiation, which will be discussed later.

1.1.5 Optimal choice when borrowing and lending rates differ

Fisher's (1930) framework assumes that the borrowing and lending interest rates in the markets are identical in a perfectly competitive capital market. However, this is unrealistic under most of the circumstances. According to the website of the Bank of England (2016), the saving interest rates are approximately between 0.5% and 1% per annum while the borrowing interest rates from credit cards are approximately between 15% and 25% per annum. Apparently, there is a wide gap between the borrowing rate and the lending rate available in the market.

Cubitt and Read (2007) analysed the situation where the borrowing and the saving interest rates were different (see also Coller & Williams, 1999). Suppose the borrowing rate is 20% per annum and the saving rate is 1% per annum. Intuitively, there would be two different interest rates to calculate the net present value when choosing between two or more options. If the calculation of net present values according to both rates favour the same option, then people should choose the option with the highest net present value. However, if the net present value according to the two interest rates favour different options, then decision makers' optimal choice depends on the financial circumstances they are facing. If decision makers are saving at the rate of 1% per annum, they should choose the option that is favoured by the net present values according the annual rate of 1%. By contrast, if they are borrowing at the rate of 20%, she should choose the option that is favoured by the net present values according to the annual rate of 20%.

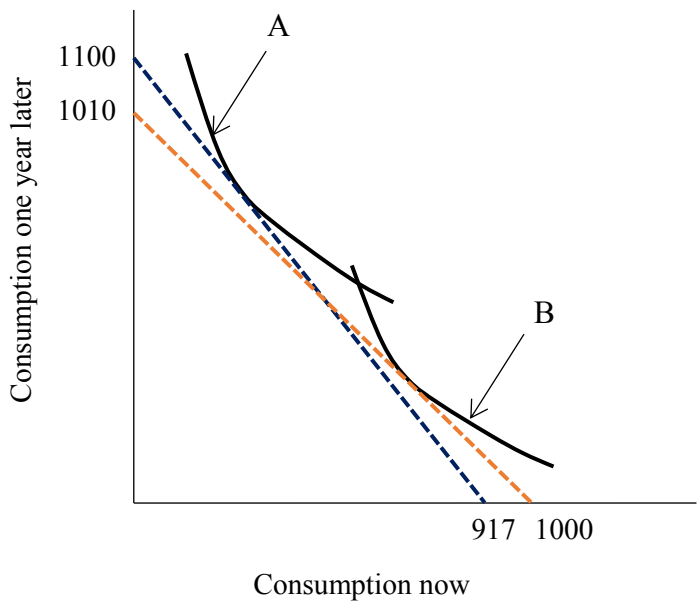

Figure 1.6. Optimal choice when borrowing and lending rates differ. The orange dashed line is the budget constraint from receiving $1000 today (SS), with the saving interest rate of 1% per annum. The blue dashed line is the budget constraint from receiving $1100, with the borrowing interest rate of 20% per annum (LL).

If the net present value according to the two interest rates favour different options and if the decision maker has neither saving nor debt, the NPV model no longer offer a conclusive answer. Consider, for example, a choice between receiving

$1000 now (referred to as SS) and receiving $1100 in one year (referred to as LL), implying an interest rate of 10% per annum. Suppose the two options are only allocated for consumption at two periods: now and in-one-year. As shown in Figure 1.6, on one extreme, if the SS option is all spent now, the amount for consumption now is $1000. On the other extreme, the SS option can be all saved with an annual rate of 1% and thus the amount for consumption in one year is $1010. It can also be partly spent now and partly saved for one-year later. So the budget constraint line from the SS option is shown as the orange dashed line in Figure 1.6. Likewise, the LL option can be all spent in one year ($1100), all spent now (borrowing $\frac{\$1100}{1+20\%}$ = $917 now for consumption and paying back $1100 in one year), or partly spent now and partly spent in one year. So the budget constraint line from the LL option is shown as the blue dashed line in Figure 1.6. Importantly the budget constraint lines offered by the two options intersect at certain point. Thus, decision makers' optimal choice will depend on their personal intertemporal preferences for consumption and should choose the option that bring them to the best indifference curve, rather than the interest rates in the market (Cubitt & Read, 2007). For example, as shown in Figure 1.6, A should choose the LL option because it is the LL option that brings her to her best attainable indifference curve while, for the same reason, B should choose the SS option.

1.1.6 Interim summary

With both the NPV model and the DU model, we can reconsider the diverse claims made in the literature. First, for tradable goods (in a perfectly competitive capital market), the NPV model, which has an interest rate related to the interest rate available in the market, should be the normative model to evaluate the choice among two or more options. With the NPV model, the interest rate elicited from laboratory experiments can be compared with the prevailing interest rate in the market. Second, the DU model, which endorses individual-specific discount rate, is the normative model for intertemporal allocation of consumption for the sake of time consistency. This interest rate is incomparable with the prevailing interest rate in the market. Third, for each individual person, the discount rate in the DU model should be the same for different goods for the sake of time consistency.

1.2 Empirical Findings

Empirical research into intertemporal choice has used diverse types of outcomes, including money, tradable commodities and directly consumable goods.

These studies, with few exceptions, have shown substantial departure from the normative predictions.

1.2.1 Monetary outcomes

Among all, many studies on intertemporal choice used monetary outcomes. The most frequently used tasks are choices between smaller-sooner (SS) and larger-later (LL) monetary options. Researchers always infer the discount rate from decision makers' choice in SS-LL questions. For example, if someone choose SS in a choice between receiving $100 today (SS) or receiving $150 in a year (LL), the choice suggests that the decision maker requires an interest rate over 50% per annum. Some also used choices between sequences of monetary outcomes. Because of the liquidity of money, it is reckoned that people can reschedule the consumption of the money by borrowing from and saving in the market. Thus the net present value (NPV) model is considered the economic standard for intertemporal choice of monetary outcomes. In this section, I review findings from the studies using monetary outcomes.

Excessive discounting. *Excessive discounting* refers to that the discount rates elicited from intertemporal choice between monetary outcomes are excessively higher than what is available in the market (Frederick et al., 2002; Read et al., 2005). As discussed earlier, the interest rate at which one can get from saving in banks is relatively low. The interest rates people can earn from other investments (e.g., securities, bonds and stocks) are generally not be too high either. The interest rates that one need to pay for loans are higher, but are mostly lower than 30% per annum in the UK or other western countries. However, in laboratory or field experiments, people always require interest rates higher than 100% per annum (see Frederick et al., 2002 for a review). For example, many people would prefer receiving $100 now to receiving $250 in one year.

Sign effect. The *sign effect* is that people discount more steeply for delayed gains than for delayed losses (Thaler, 1981). For example, someone indifferent between receiving $100 today and receiving $150 in one year would probably prefer paying $100 today to $150 in one year. The sign effect is replicated in many studies (e.g., Estle et al., 2006; Ohmura, Takahashi, & Kitamura, 2005; Xu, Liang, Wang, Li, & Jiang, 2009). In some case, the sign effect can be more extreme that many people will prefer to pay now rather than to pay later, even if delaying the payment does not incur additional financial costs (e.g., Hardisty, Appelt & Weber, 2013; Yates & Watts, 1975).

Absolute magnitude effect. The absolute magnitude effect is that implied discount rates decrease with the magnitude of the outcome (Loewenstein and Prelec 1992). For example, someone indifferent between $200 in one year and $100 today would probably prefer $2,000 in one year to $1,000 today. It is one of the most robust phenomena in intertemporal choice and has been corroborated in a wide range of studies (e.g., Baker, Johnson, & Bickel, 2003; Benhabib, Bisin, & Schotter, 2010; Benzion, Rapoport, & Yagil, 1989; Green, Fristoe, & Myerson, 1994; Holcomb & Nelson, 1992; Petry, 2001; Thaler, 1981).

Delay effect. The delay effect is that people require higher interest rate for short delays than for long delays (Thaler, 1981). For example, someone indifferent between $100 today and $225 in two years (implying an interest rate of 50% per annum) would probably prefer $100 today to $150 in one year (implying an interest rate over 50% per annum). Concerning the delay effect, there are two different explanations. One is decreasing impatience and the other is subadditive discounting (Read, 2001). See further discussion of the two effects below.

Non-constant discounting. Non-constant discounting includes decreasing impatience and increasing impatience. In some literature, *decreasing impatience* is also known as the common difference effect, which means that the interest rate over an interval decreases as the delay to the onset of the interval increases (Loewenstein & Prelec, 1992). For example, someone indifferent between $200 in one year and $100 today would probably prefer $200 in two years to $100 in one year. This effect has been corroborated in several studies (e.g., Green et al., 1994; Green, Myerson, & Macaux, 2005; Holt Green, Myerson, & Estle, 2008; Keren & Roelofsma, 1995; Kirby & Herrnstein, 1995; Scholten & Read, 2006). However, some others have found evidence for *increasing impatience* that the discount rate over an interval increase with the onset of the interval (e.g. Holcomb & Nelson, 1992; Read et al., 2005; Sayman & Öncüler, 2009; Attema et al., 2010).

Non-additive discounting. Non-additive discounting is that the discounting over an interval depends on whether it is divided into sub-intervals or is kept undivided. Two patterns of non-additivity have been observed: subadditivity and superadditivity. *Subadditivity* is that implied discount rates are lower over an undivided interval than over an interval that is divided into subintervals. It has been corroborated in many studies (Kinari, Ohtake, & Tsutsui, 2009; McAlvanah, 2010; Read, 2001; Read & Roelofsma, 2003; Scholten & Read 2006). An example for subadditivity is that

someone indifferent between $100 now and $150 in six months and indifferent between $150 in six months and $200 in one year would prefer $200 in one year to $100 today. *Superadditivity* is the reversal of subadditivity, in that implied discount rates are higher over an undivided interval than over an interval that is divided into subintervals (Scholten & Read, 2006; 2010; Scholten et al., 2014).

Sequence effects. Intertemporal choices between sequences of positive outcomes have been shown to be different from those between single-dated outcomes. Most compellingly, *negative time preference*, which means a preference for a positive outcome to take place later rather than earlier, has been frequently observed when participants choose between intertemporal sequences (Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991; Read & Powell, 2002), but is rarely observed when they are asked to choose between two single-dated outcomes.

Framing effects. The literature has documented diverse framing effects showing how the description of intertemporal choice questions influences intertemporal choice. The *delay/speed-up asymmetry* is that people are less impatient when the intertemporal choice is framed as expediting the larger-later option to the smaller-soon option than when the same choice is described as deferring the smaller-soon option to the larger-later option (Appelt, Hardisty, & Weber, 2011; Benzion et al., 1989; Loewenstein, 1988; Malkoc & Zauberman, 2006; McAlvanah, 2010; Scholten & Read, 2013; Shelley, 1993; Weber, Johnson, Milch, Chang, Brodscholl, & Goldstein, 2007). The *date/delay effect* is that people tend to be less impatient when time is described with calendar dates than with the length of delays (LeBoeuf, 2006; Read et al., 2005b). *Outcome framing effects* are that people exhibit different degrees of impatience when the outcomes are framed in different terms, such as an interest rate, the gross interest earned, or the total amount earned (Read et al., 2005a; Read, Frederick, & Scholten, 2013). The *(asymmetric) hidden-zero effect* is that the explicit display of getting nothing at a later point of time in the SS option reduces impatience (Magen, Dweck, & Gross, 2008; Read et al., in press; Wu & He, 2012). Lastly, Comparisons of different *response modes* always show that time preference elicited with different methods are systematically different (Hardisty, Thompson, Krantz, & Weber, 2013; Read & Roelofsma, 2003).

1.2.2 Non-monetary outcomes

Although a majority of studies used monetary outcomes for research into intertemporal choice, there is still a significant proportion of studies using other commodities as outcomes, such as food, drinks and other consumable goods.

Domain-specific discounting. Most of the studies on the discounting of consumable commodities made a comparison between the interest rates of consumable commodities and that of monetary outcomes. The results often suggested that directly consumable goods (e.g., alcohol, food, CDs and DVDs) are discounted more steeply than money (e.g., Charlton & Fantino, 2008; Estle et al., 2007; Odum, Baumann, & Rimington, 2006; Odum & Rainaud, 2003; Tsukayama & Duckworth, 2010; Ubfal, 2016). There are some exceptions though. Hardisty and Weber (2009) found that the discount of monetary outcomes, (public) environmental goods and health was similar to each other. Chapman (1996b) found that the discount rate for health was always higher than that for money.

While there is always a gap between discounting of money and consumable commodities, individuals' discount of money and consumable commodities are still correlated. Reuben, Sapienza and Zingales (2010) found that the discount rates for money and chocolates were moderately correlated. Odum (2011) showed that the discount rate for money was highly correlated with the discount rates for a variety of consumable commodities (i.e., food, heroin and cigarettes). Ubfal (2016) also found that high degrees of correlation among discount rates of a variety of goods including money, meat and sugar. Tsukayama and Duckworth (2010) showed that the correlation between discount rates of money and that of consumable goods were lower than the correlation of discount rates among different consumable goods, drawing a second line between directly consumable goods and money.

A surprising observation in these studies mentioned above is that when the discount rates of different goods were compared, none of them used the DU model to estimate individual discount rates. Instead, discount rates are usually measured with the quasi-hyperbolic discounting model (Laibson, 1997), the hyperbolic discount model (Mazur, 1987) and/or a model-free estimation called Area Under the Curve (AUC: Myerson, Green, & Warusawitharana, 2001).

Utility function. Although many studies claimed to identify domain-specific discount rates, the vast majority failed to address the curvature of the utility function properly. Many of them made an assumption that the utility from consumption is a

linear function of amount of consumption, even for food (see Kirby and Santiesteban, 2003 for an exception). A closely related issue to the utility function is satiation of consumptions (Read et al., in press). For example, someone having two pizzas for oneself to eat will not be twice as happy as when she has only one pizza to eat. So, an important empirical concern is the elicitation of utility function.

Measuring the curvature of the utility function is a difficulty. Different approaches have been proposed. Andersen et al., (2008) used the method of double elicitation. With monetary outcomes, Andersen et al. (2008) elicited the curvature of the utility function for money from risky choice and applied the utility function from risky choice to intertemporal choice. Some others have attempted to elicit time preference by avoiding the curvature of the utility function through their experimental designs (Attema et al., 2010; Chapman, 1996b; Laury, McInnes, & Swarthout, 2012). Although the technical solutions of these designs differ from one another, they share the same intuition that the amounts of the outcomes in SS and LL are kept constant across items. However, these methods were applied to monetary outcomes only. Further research should pay more attention to the curvature of the utility function.

1.3 Causes of Departures from the Normative Models

Empirical tests of intertemporal choice of money or non-monetary goods have shown substantial departures from the predictions by the NPV model or the DU model. Various causes have been proposed to explain the departures including both environmental factors and cognitive limitations.

1.3.1 Market imperfection and complication

The NPV model makes a strong assumption of a perfect capital market where people have perfect and costless access to saving and borrowing opportunities. However, this is by no means true for decision makers. First of all, as stated earlier, there is a large gap between the rates with which decision makers can save and borrow. Second, the interest rate available in the market depends on the amount of money saved or borrowed. For example, there is probably no such a formal market for one to exchange £1 now for £1.05 in one year, but it is likely that one can exchange £1,000 now for £1,050 in one year. Third, transaction costs are inevitable when people reschedule their consumption across time by lending and borrowing. The optimal choice is not cognitively straightforward, especially when the borrowing and lending rates available to a person differ (Cubitt & Read, 2007). In addition, people have to at least go to ATM or logon to online bank account to transfer money.

Concerning the variability of discount rates for different commodities, Keynes (1936) points out that the exchange rate between commodities is not constant across time. Indeed, even the exchange rate between two currencies, such as sterling pounds and dollars, keeps changing over time. Other important factors that influence commodity-specific interest rates include the yield, the carrying cost and the liquidity premium (Keynes, 1936; Read et al., in press). The *yield* means the owner's benefit from the use of a good, especially when the value of the good is not much reduced after the use. The yield is especially applicable to durable goods, such as a house. The owner of the house can use the house as accommodation, but the value of the house in the market is more or less maintained. The *carrying cost* refers to the cost incurred during the good is held, such as storage, obsolescence, spoilage and wastage. A good example is fresh vegetables sold in supermarkets. The *liquidity premium* refers to the convenience of a good to be exchangeable to other goods with the identical market value. Among all, money, as the common currency, probably has the highest liquidity premium. For example, one can buy candies or shop groceries with money. But it is inconvenient to exchange candies for groceries or vice versa.

1.3.2 Background consumption

While background consumption is one of most important determinants of the observed time preference in Fisher's (1930) framework, it often dismissed in the vast number of studies investigating individual differences in intertemporal preference (e.g., Dohmen, et al., 2010; Enzler et al., 2014; Shamosh et al., 2008) and discounting of different goods (e.g., Augenblick et al., 2012; Chapman, 1996b; Estle et al., 2007). It is especially problematic when the curvature of the utility function (in the DU model) is not taken into consideration. For example, Noor (2009) used the curvature of the utility function induced by background consumption to explain hyperbolic discounting on raw amounts.

Moreover, it is almost indistinguishable between the utility function and the discount rate in the DU model (Read et al., in press). For example, when someone is highly desired for an immediate consumption at the cost of the consumption later of a much larger amount, it could be attributed to a high discount rate but it could be equivalently attributed to a disproportionately large utility from the current consumption.

Fisher's (1930) definition of the pure time preference, which appears a good solution that sidesteps the difficulty of the elicitation of the curvature of the utility

function. However, as discussed earlier, this pure time preference is defined as the marginal rate of intertemporal substitution *only* when the background consumption is equal at different time periods. Thus it is often unobservable and is unlikely to overcome the obstacles from background consumption either.

1.3.3 Uncertainty associated with delayed offers

Many people may be concerned with the uncertainty of the delayed outcome when faced with a choice between an immediately available and a delayed outcome (Epper, Fehr-Duda, & Bruhin, 2011; Fisher, 1930). Sozou (1998) shows that a hyperbolic discount function can arise from the (constant) exponential discounting function if future outcomes are uncertainty. Michaelson et al. (2013) showed that people are more likely to choose the delayed but larger reward when the person who is going to deliver it looks trustworthy than when the person looks untrustworthy, which implies that the delayed option is perceived as risky in the meantime.

Because of the confound of uncertainty, some studies took measures to control the uncertainty of a delayed payment. For example, Andreoni and Sprenger (2012) tried to guarantee the delayed payments using a credit system. In Andersen, Harrison, Lau and Rutstrom (2008), delayed payments to participants were guaranteed by a national Ministry in Denmark and was paid directly into participants' personal bank accounts.

1.3.4 Subjective perception of delays

Psychophysical accounts suggest that subjective perception of or sensitivity to delays is a source of non-constant discounting. Takahashi (2005) shows that a generalized hyperbolic discount function by Loewenstein and Prelec (1992), which embodies decreasing impatience, can be derived from constant impatience with a decreasingly elastic function for subjective time perception. Experimental studies have supported this view by showing that the subjective perception of delays is nonlinear and that the discounting over the subjective perception of delays tends to be constant (Han &Takahashi, 2012; Zauberman, Kim, Malkoc, & Bettman, 2009). Ebert and Prelec (2007) extended the view of subjective sensitivity to delays to allow for both decreasing and increasing impatience.

1.3.5 Myopia or a lack of self-control

Explanations to time-inconsistent intertemporal choice always pinpoint to the problem of myopia or self-control (Ainslie, 1975; Ainslie & Herrnstein, 1981; Strotz, 1955). Myopic individuals often give special weights to temporally proximal selves

and over-consume the resources and thus leave less for future selves. When future selves come closer in time, the "future" selves become the "present" ones and again over-consume. When this iterative process happens for multiple selves, the far-future selves will get almost nothing (Figure 1.4). This lays the foundations of many studies of intertemporal choice in behavioural economics and has been applied to diverse phenomena such as pension scheme (Laibson,1997), credit card borrowing (Meier & Sprenger, 2012), procrastination (O'Donoghue & Rabin, 1999) and addiction (Heyman, 1996).

1.3.6 Cognitive inertia

Lastly, but probably most importantly for the thesis, cognitive inertia could be a key drive behind intertemporal choice. As shown by Frederick (2005), a simple Cognitive Reflection Test (CRT) consisting of three items is a prominent predictor of intertemporal discounting. CRT is a psychological battery that assesses to what degree one is likely to use intuitive heuristics or deliberative thinking, analogous to dual-systems (Sloman, 1996; Kahneman and Frederick, 2002). Based on CRT, those with high cognitive reflection are much less impatient than those with low cognitive reflection in intertemporal choice of monetary outcomes.

Various framing effects are also strong evidence for the key role of cognitive inertia in intertemporal choice. These effects suggest that most decision makers not only skip optimising, but also make decisions in respect to many normatively irrelevant information (see *framing effects* in Section 1.2.1, pp. 17). A good example is from Read et al. (2013), their participants exhibited much less impatience when the intertemporal choice between monetary outcomes was described as an investment, while the normative account suggest that people are always aware of investment opportunities. It is these framing effects that call for a more coherent understanding of the psychology of intertemporal choice.

1.4 Cognitive Processes in Intertemporal Choice

The failure of the normative models has motivated some researchers to move away from the normative approach and to study what cognitive processes underlie intertemporal choice, in line with the recommendation by Ariel Rubinstein (2003) to open the psychological black-box of intertemporal decision making. The reasons for this shift are probably twofold. First, the reality in the environment does not satisfy the assumptions in the normative models. For example, for the NPV model, a perfect competitive capital market hardly exists. When the model fails to capture realistic

environmental features relating to the making of decisions, even rational individual behaviour could depart from the prediction of the model. Second, various framing effects suggest that individuals show inconsistent preferences when normatively equivalent decisions (see *framing effects* in Section 1.2.1, pp. 17). Those effects are frequently explained by different cognitive processes invoked by different framing, highlighting the importance of studying cognitive processes in intertemporal choice (e.g., Cubitt, McDonald, & Read, 2017; Read et al., 2016; Weber et al., 2007). Below, I briefly introduce several important cognitive processes that could govern individual intertemporal choice, which are also the focus of the following chapters in the thesis.

1.4.1 Evaluation rule

One of the fundamental questions regarding the cognitive processes of intertemporal choice, as well as other types of value-based decision making, is how information of the choice is processed and evaluated. There are mainly two different evaluation rules in intertemporal choice: alternative-based and attribute-based rules (see Scholten et al., 2014; Dai & Busemeyer, 2014).[9] According to the alternative-based rule, options are evaluated and assigned values independently. The option with the highest assigned value is chosen. By contrast, the attribute-based rule assumes that options are directly compared along the time and outcome attributes respectively, and the option favoured by these comparisons is chosen (e.g., Gonzalez-Vallejo, 2002; Tversky, 1972).

Two strands of evidence could shed light on the comparisons of different evaluation rules. First, eye-tracking studies offer the process-level evidence on whether decision is made by alternative-wise computation or attribute-wise comparison. For example, Arieli, Ben-Ami and Rubinstein (2011) found that participants make intra-attribute eye movements much more often than intra-option eye movements. Second, different evaluation rules are written in models and thus the descriptive accuracy of the models with different evaluation rules behind can be quantitatively compared with each other (e.g., Scholten et al., 2014; Dai & Busemeyer, 2014).

[9] A third evaluation rule was used in Scholten and Read's (2006) interval discounting model. See further discussion in Chapter 2.

1.4.2 Attention

Attention allocation matters in many value-based decision making, including intertemporal choice. Many studies posit that attention is a key cognitive process that drives value-based preference (e.g., Bhatia, 2014; Bordalo, Gennaioli, & Shleifer, 2012; Kőszegi & Szeidl, 2013; Read et al., 2013; Tsetsos, Chater, & Usher, 2012; Tversky, Sattath, & Slovic, 1988; see Weber & Johnson, 2009 for a review). A few studies have directly tested the relationship between attention and food choice (Armel, Beaumel, & Rangel, 2008; Krajbich, Armel, & Rangel, 2010), risky choice (e.g., Fiedler & Glőckner, 2012; Stewart, Hermens, & Matthews, 2016) and intertemporal choice (e.g., Fisher & Rangel, 2014; Franco-Watkins, Mattson, & Jackson, 2016). A better understand of the attention effects on intertemporal choice is a key step towards the understanding of cognitive processes of intertemporal choice.

1.4.3 Background contrast

Background contrast is another important cognitive factor that have been found to play an important role in value-based decision making (Simonson & Tversky, 1992). This effect has been very well-established in a wide range of studies (Ebert & Prelec, 2007; Priester, Dholakia, & Fleming, 2004; Simonson & Tversky, 1992; Vlaev & Chater, 2006; 2007; Walasek & Stewart, 2015). In their seminal work, Simonson and Tversky (1992) showed that participants' tradeoff between two attributes (e.g., the price and the quality of a personal computer) was strongly influenced by the preceding tradeoff they have been exposed to, regardless of their choices in the preceding tradeoff.[10] For example, as shown in Figure 1.7, in a choice between paying $1200 for a computer with 720K memory (Option *x*) or paying $1000 for a computer with 640K memory (Option *y*), participants were more likely to choose *x* when they had been previously exposed to a choice between paying $1200 for a computer with 880K memory (Option *a*) or paying $1000 for a computer with 480K memory (Option *b*) than when they had been previously exposed to a choice between paying $1500 for a computer with 720K memory (Option *a'*) or paying $700 for a computer with 640K memory (Option *b'*).

[10] A closely related effect to the background contrast effect is the prospect relativity effect (e.g., Stewart, Chater, Stott, & Reimers, 2003; Ungemach, Stewart, & Reimers, 2011; Vlaev, Chater, & Stewart, 2007; Vlaev, Seymour, Dolan, & Chater, 2009), which suggests that the valuation of an option (or an attribute value), rather the tradeoff rate between two attributes, is influenced by previously exposed similar options (or values of the same attribute).

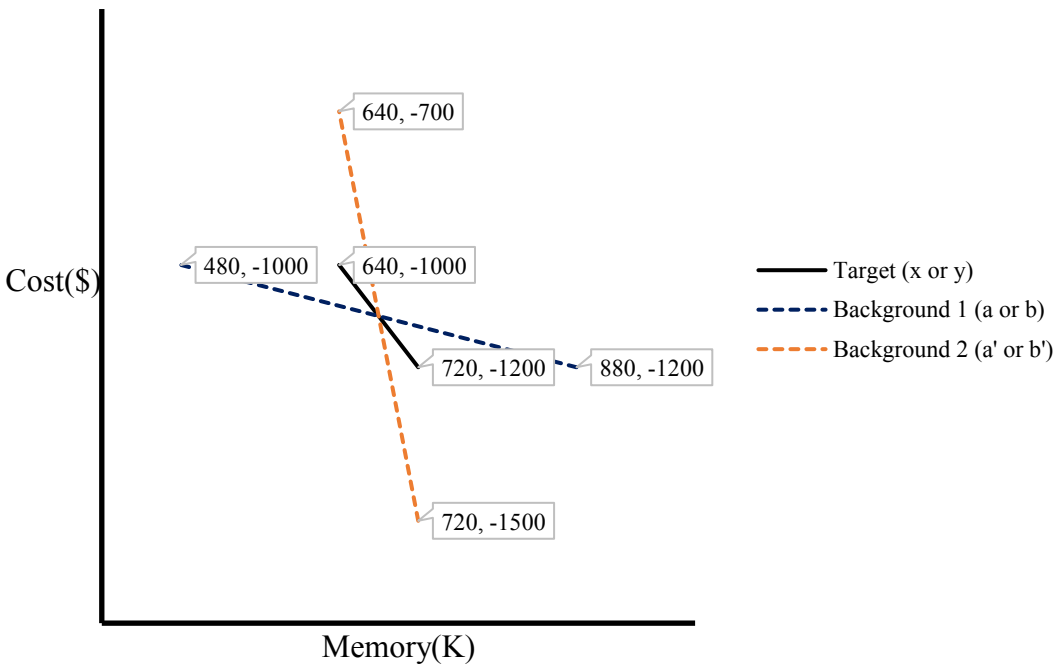

Figure 1.7. A graphical illustration of Simonson and Tversky's (1992) study on the background contrast effect.

1.5 Overview of the Following Chapters

The thesis is aimed at a better understanding of the cognitive processes in intertemporal choice. Three separate empirical studies are reported in Chapters 2-4, investigating various cognitive processes including evaluation rules, attention, and background contrast. Chapter 5 further summarises the results from the empirical studies and discusses the implications on theory development of intertemporal choice.

The study in Chapter 2 makes a comprehensive model comparison in intertemporal choice. The literature has documented a wide array of intertemporal choice models. However, existing studies that quantitatively compare intertemporal choice models often lend support to different models. Several limitations make it difficult to compare the results from different studies. This comprehensive model comparison attempts to address or, at least, alleviate three limitations: Model selectivity, stochastic-specification selectivity and stimulus diversity. The results lend strong support to attribute-based models, especially the family of the tradeoff model. There is also an interaction between models and stochastic specifications on model performance, highlighting the importance of stochastic specification in intertemporal choice modelling.

The study in Chapter 3 investigates the attention effect on intertemporal choice. Previous studies on the relationship between attention and value-based choice often assumed an option-wise attention effect and other ways of the attention effect. The study in Chapter 3 develops a new attention manipulation method that allows the attention effect to operate in multiple ways to test the attention effect on intertemporal choice. The results from the study suggests that, beyond the option-wise attention effect, attention can operate in attribute-wise and component-wise ways to influence intertemporal choice. In addition, the study observes robust background contrast effects on intertemporal choice, which indicate the violation of sequential independence in intertemporal choice.

The study in Chapter 4 focuses on a controversial topic in intertemporal choice: the patterns of impatience. Concerning the degree of impatience in near vs. far future, people can show decreasing impatience, constant impatience and increasing impatience. As the basis of the family of hyperbolic discounting models, decreasing impatience has received much attention these years. However, testing of the patterns impatience has accumulated mixed and conflicting results. The study in Chapter 4 tests two ways to look through the heterogeneity. The first is on a potential design bias. A vast number of tests of the pattern of impatience used pairs of intertemporal choice items. This method makes an implicit assumption that individuals' patterns of impatience are independent from their degrees of impatience. The first goal of the study is to directly test the hypothesis of independence between the pattern and the degree of impatience. Results suggest that the two are correlated and thus point to a design bias when pairs of items are used to test the aggregate pattern of impatience. The second focus is on the instability of the detected pattern of impatience from experimental studies. Specifically, an order effect, which is relevant to the background contrast effect, on the detected aggregate pattern of impatience is observed. Taken together, the two ways may reconcile much of the variation in the detected aggregate pattern of impatience in the literature.

Chapter 5 provides a general discussion of the thesis. First, I discuss the results from Chapters 2-4 in the general framework of the cognitive processes in intertemporal choice, crystallising the understanding of them. It further points out the limitation of the thesis as it only investigated one type of intertemporal choice and potential extensions to other domains of intertemporal choice are discussed.

CHAPTER 2 A COMPREHENSIVE MODEL COMPARISON

Experimental studies on intertemporal choice typically investigate choices between smaller-sooner (SS) and larger-later (LL) amounts of money, such as a choice between $100 today and $200 in one year. Economic analysis provides researchers and consumers with "normative" models of how people should make these choices, but as with many other domains of choice, economic models do not predict the preferences people express. Consequently, researchers have developed alternative accounts of how intertemporal choices for money are made, each designed to capture the complex interaction between the choice options people are presented with, and their preferences over those options (e.g., Doyle, 2012).

In this study, we systematically addressed the question of which models and categories of models best described human choice behaviour. We did this by means of quantitative model comparisons, in which models were tested against empirical data. We report a comprehensive quantitative comparison between models of intertemporal choice drawn from economics and psychology. We addressed major limitations in previous model contests by including fifteen intertemporal choice models, three stochastic specifications and 225 (out of 256) data sets collected from diverse sources.

2.1 Categories of Models

The simple intertemporal choices we investigated involved two options. The SS option is a smaller amount of money (denoted x_S) at an earlier time (x_L). The LL option is a larger amount (x_L) at a later time (t_L). Choices between such options can be determined by a range of transformations and interactions of these four values. One possibility is what we call alternative-based choice. Alternative-based choice means that each alternative (either SS or LL) is independently assigned a value, and these values are then compared with the better option being chosen. In intertemporal choice, alternative-based choice always entails some function of time (denoted $d(t)$ in Table 2.1) being used to weight some function of the outcome (denoted $v(x)$), so the present value of a delayed outcome is $v(x)d(t)$.

Table 2.1 Categories of intertemporal choice models.

Category	Evaluation rule
Alternative-based	$v(x_S)d(t_S) \gtreqless v(x_L)d(t_L)$
Attribute-based	$Q(t_S, t_L) \gtreqless V(x_S, x_L)$
Hybrid	$D(0, t_S)v(x_S) \gtreqless D(0, t_S)D(t_S, t_L)v(x_L)$

Attribute-based models are derived from an approach to choice initiated by Tversky's (1969) additive difference model. In these models, time and outcome are treated as separate attributes just as in other multi-attribute choice models, and the advantage or disadvantage of each option on each attribute is computed, and the choice is made by evaluating whether a sum of attribute differences favours one option or another. When choosing between \$100 now and \$200 in one year, for instance, the decision maker determines how much better \$200 is from \$100 (computing $V(x_S, x_L)$), and how much better "now" is to "one year" (computing $Q(t_S, t_L)$) and chooses the \$200 if the outcome advantage exceeds the time advantage, and the \$100 otherwise. Attribute-based models were pioneered by Leland (2002) and Rubinstein (2003), and a mathematical model amenable to formal testing was described by Scholten and Read (2010).

Hybrid models mix alternative-based and attribute-based comparisons. Only one hybrid model is in the literature, the interval discounting model (Read 2001, Scholten and Read 2006, Scholten et al. 2014). This model is alternative-based in that each outcome receives a discounted value, and it is attribute-based in that the values of outcomes are discounted not only as a function of the delays to the outcomes, but also as a function of the interval between them.

2.2 Models

The primary goal of this study was to compare different categories of models. The fifteen intertemporal choice models listed in Table 2.2 were subjected to a model comparison. These include eight alternative-based models, one hybrid model, and six attribute-based model. Seven of the eight alternative-based models are what we will call *outcome discounting models*, because they do not involve any transformation of monetary outcomes, and thus only differ from the standard economic model in the

way discounting takes place. These models are (1) the *exponential* discounting model, (2) the *quasi-hyperbolic* discounting model (Laibson, 1997), which includes (1) as a special case, (3) a *hyperbolic* discounting model (Herrnstein, 1981; referred to as $\text{Hyperbolic}_{\text{H}}$), (4) a *generalized hyperbolic* discounting model (Loewenstein & Prelec, 1992; referred to as $\text{Hyperbolic}_{\text{LP}}$), which includes (1) and (3) as special cases, (5) a hyperbolic discounting model in which discounting takes place over weighted delays (Mazur, 1987; referred to as $\text{Hyperbolic}_{\text{M}}$), (6) the *constant-sensitivity* discounting model, which includes (1) as a special case, and in which, depending on how delays are weighted, discounting can be hyperbolic *or* hypobolic (Ebert & Prelec, 2007), and (7) the *double-exponential* discounting model (McClure et al., 2007). The eighth alternative-based model, Loewenstein and Prelec's (1992) full hyperbolic discounting model, is a value-discounting model, in which the generalized hyperbolic discount function is applied to subjective outcome values rather than objective outcomes. The specific operationalization used here comes from Scholten, Read and Sanborn (2014), thus it is referred to as $\text{Hyperbolic}_{\text{SRS}}$.

The hybrid model is the *discounting by intervals model* (Interval, Read, 2001; Scholten & Read 2006; Scholten et al., 2014). Being a generalization of $\text{Hyperbolic}_{\text{SRS}}$, it is also a value-discounting model, and with an increasingly elastic value function.

The six attribute-based models involved in the model contest are the *proportional-difference model* (PD; Cheng & González-Vallejo, 2016), the *intertemporal-choice-heuristics model* (ITCH; Ericson et al., 2015), the *difference-ratio-interest-finance-time model* (DRIFT; Read et al., 2013), the *tradeoff model* (TM; Scholten et al., 2014), and two reduced versions of the tradeoff model which we call *basic tradeoff models*. They differ in whether the value function and the time-weighing function are normalized logarithmic functions (Scholten & Read, 2013; referred to as BTM_{SR}, a special case of TM) or power functions (Dai & Busemeyer, 2014; referred to as BTM_{DB}).[11]

[11]The tradeoff model evaluated by Scholten and Read (2013) reduces to BTM_{SR} when the data do not contain the delay-speedup asymmetry, as is the case in our data sets.

Table 2.2 List of intertemporal choice models involved in this model comparison.

ALTERNATIVE-BASED MODELS				
Name	Source	Delay Discount Function	Value Function	Domain
Exponential	Samuelson (1937)	$d(t) = \delta^t$	$v(x) = x$	$0 \le \delta \le 1$
Quasi-hyperbolic	Laibson (1997)	$d(t) = \beta^{\mathrm{I}(t>0)}\delta^t$	$v(x) = x$	$0 \le \beta,\ \delta \le 1$
Hyperbolic$_{\mathrm{H}}$	Herrnstein (1981)	$d(t) = (1+\alpha t)^{-1}$	$v(x) = x$	$\alpha \ge 0$
Hyperbolic$_{\mathrm{M}}$	Mazur (1987)	$d(t) = (1+\alpha t^{\tau})^{-1}$	$v(x) = x$	$\alpha,\ \tau \ge 0$
Hyperbolic$_{\mathrm{LP}}$	Loewenstein and Prelec (1992)	$d(t) = (1+\alpha t)^{-\beta/\alpha}$	$v(x) = x$	$\alpha,\ \beta \ge 0$
Constant sensitivity	Ebert and Prelec (2007)	$d(t) = e^{-(\beta t)^{\tau}}$	$v(x) = x$	$\beta,\ \tau \ge 0$
Double exponential	McClure et al. (2007)	$d(t) = \omega\beta^t + (1-\omega)\delta^t$	$v(x) = x$	$0 \le \beta,\ \delta \le 1,\ 0.5 \le \omega \le 1$
Hyperbolic$_{\mathrm{SRS}}$	Scholten et al. (2014)	$d(t) = (1+\alpha t)^{-\beta/\alpha}$	$v(x) = (1-\gamma)x^{(1-\gamma)} + \mu\gamma x^{\gamma}$	$\alpha,\ \beta \ge 0,\ 0 \le \gamma \le 1,\ \mu \ge 1$
ATTRIBUTE-BASED MODELS				
Name	Source	Time and Outcome Advantages		Domain
PD (Logistic / Fechnerian only)	Cheng and González-Vallejo (2016)	$V(x_S, x_L) = \dfrac{x_L - x_S}{x_L}$ $Q(t_S, t_L) = \dfrac{t_L - t_S}{t_L} + \psi$		$-2 \le \psi \le 2$
PD (Luce only)[a]		$V(x_S, x_L) = \dfrac{x_L - x_S}{x_L}$ $Q(t_S, t_L) = \kappa\dfrac{t_L - t_S}{t_L}$		$\kappa \ge 0$
ITCH[b]	Ericson et al. (2015)	$V(x_S, x_L) = (1-w_x)(x_L - x_S) + w_x\dfrac{x_L - x_S}{x_*}$		$0 \le w_x \le 1$

DRIFT	Read et al. (2013)	$Q(t_S,t_L)=\kappa\left((1-w_t)(t_L-t_S)+w_t\,\frac{t_L-t_S}{t_*}\right)$	$0\le w_t\le 1,\ \kappa\ge 0$
		$V(x_S,x_L)=(1-w_I)(1-w_R)(x_L-x_S)+w_R\left((x_L-x_S)/x_S\right)+w_I\left((x_L/x_S)^{(1/(t_L-t_S))}-1\right)$	$0\le w_I,\ w_R\le 1$
		$Q(t_S,t_L)=\kappa(t_L-t_S)$	$\kappa\ge 0$
BTM_{DB}	Dai and Busemeyer (2014)	$V(x_S,x_L)=x_L^{\ \gamma}-x_S^{\ \gamma}$	$0\le\gamma\le 1$
		$Q(t_S,t_L)=\kappa(t_L^{\ \tau}-t_S^{\ \tau})$	$\kappa\ge 0,\ 0\le\tau\le 1$
BTM_{SR}	Scholten and Read (2013)	$V(x_S,x_L)=\frac{1}{\gamma}\left(\ln(1+\gamma x_L)-\ln(1+\gamma x_S)\right)$	$\gamma\ge 0$
		$Q(t_S,t_L)=\frac{\kappa}{\tau}\left(\ln(1+\tau t_L)-\ln(1+\tau t_S)\right)$	$\kappa,\ \tau\ge 0$
TM	Scholten et al. (2014)	$V(x_S,x_L)=\frac{1}{\gamma}\left(\ln(1+\gamma x_L)-\ln(1+\gamma x_S)\right)$	$\gamma\ge 0$
		$Q(t_S,t_L)=\frac{\kappa}{\alpha}\log\left(1+\alpha\left(\frac{(1/\tau)\left(\ln(1+\tau t_L)-\ln(1+\tau t_S)\right)}{\vartheta}\right)^{\vartheta}\right)$	$\alpha,\ \kappa,\ \tau\ge 0,\ \vartheta\ge 1$

HYBRID MODEL

Name	Source	Interval Discount Function and Value Function	Domain
Interval	Scholten et al. (2014)	$D(t_S,t_L)=\left(1+\alpha\left(\frac{(1/\tau)\left(\ln(1+\tau t_L)-\ln(1+\tau t_S)\right)}{\vartheta}\right)^{\vartheta}\right)^{-\beta/\alpha}$	$\alpha,\ \beta,\ \tau\ge 0,\ \vartheta\ge 1$
		$v(x)=(1-\gamma)x^{(1-\gamma)}+\mu\gamma x^{\gamma}$	$0\le\gamma\le 1,\ \mu\ge 1$

[a] Under the Luce specification, PD could not be estimated in its original form, because ψ could move the time advantage of SS into negative territory.

[b] x_* is the arithmetic mean of the two outcomes: $x_*=\frac{x_S+x_L}{2}$. t_* is the arithmetic mean of the two delays: $t_*=\frac{t_S+t_L}{2}$. The intercept, a free parameter, of the original ITCH model (Ericson et al., 2015) is removed. That intercept is theoretically incorrect because it will predict a preference towards a specific option in a binary choice even when the two options are identical.

Table 2.3 Behavioural regularities covered by intertemporal choice models.

Category	Model	Phenomenon						
		Positive discounting	Absolute magnitude effect	Delay effect	Common difference effect (CDE)	Reverse CDE	Nonadditivity	Relative nonadditivity
Alternative-based	Exponential	✓						
	Quasi-hyperbolic	✓		✓	?			
	Hyperbolic_{H}	✓		✓	✓			
	Hyperbolic_{M}	✓		✓	✓			
	Hyperbolic_{LP}	✓		✓	✓			
	Constant sensitivity	✓		✓	✓	✓		
	Double exponential	✓		✓	✓			
	Hyperbolic_{SRS}	✓	✓	✓	✓			
Hybrid	Interval	✓	✓	✓	✓		✓	
Attribute-based	PD	?		✓	✓			
	DRIFT	✓	✓	✓	✓			?
	ITCH	✓	✓	✓	✓			?
	BTM_{DB}	✓	✓	✓	✓			?
	BTM_{SR}	✓	✓	✓	✓			?
	TM	✓	✓	✓	✓		✓	?

Note. A check mark means that a model can accommodate the phenomenon. A question mark means that a model can accommodate the phenomenon on some occasions, but not on others. Models accounting for relative nonadditivity do so by combining with the Luce specification. Although the ITCH model can occasionally accommodate nonadditivity according to the results from Parameter Space Partitioning in Appendix 2C, however, this property is irregular and intractable, in a stark contrast to TM, which uses two parameters to systematically accommodate subadditivity and superadditivity respectively (Scholten et al., 2014).

2.3 Behavioural Regularities and Model Coverage

Studies involving choices between smaller-sooner (SS) and larger-later (LL) monetary options have discovered numerous patterns or what we call behavioural regularities. These behavioural regularities are the primary motive for developing alternative models of intertemporal choice and, on the other hand, constitute qualitative evidence for or against models (summarised in Table 2.3). Here we provide a brief description of the behavioural regularities reported for monetary gains.

2.3.1 Positive discounting

Positive discounting is that people are generally impatient and prefer a positive-valence outcome to happen sooner, rather than later (Frederick et al., 2002).[12] It is normally the cornerstone of studies employing SS-LL choices of monetary gains. Positive discounting implies that, when the smaller outcome (x_S) and the smaller delay (t_S) are held constant, an increase in the larger delay (t_L) requires an increase in x_L for a person to remain indifferent between SS and LL (Scholten & Read, 2013). For instance, someone indifferent between \$100 today and \$200 in one year would be indifferent between \$100 today and *more than* \$200 in two years. All the fifteen models can fully accommodate this except for the PD model. Specifically, the function for time advantage in PD is insensitive to t_L when $t_S = 0$, because $(t_L - 0) / t_L = 1$, i.e., the proportional difference between the delays is 1 regardless of how long t_L is.

2.3.2 Absolute magnitude effect

The absolute magnitude effect is that the discount rate decreases with outcome magnitude (Loewenstein & Prelec, 1992). For example, someone indifferent between \$200 in one year and \$100 today is likely to prefer \$2,000 in one year to \$1,000 today, when, in both choices, the interest rates on offer is a 100% per year. The absolute magnitude effect is a robust finding for monetary gains.[13] None of the outcome-discounting models can accommodate this effect but the two value-discounting models (i.e., $\text{Hyperbolic}_{\text{SRS}}$ and Interval) can accommodate the absolute magnitude effect with the increasingly elastic value function. All attribute-based models, but PD, can accommodate the absolute magnitude effect.

[12] Positive discounting is not necessarily anomalous to the NPV model, but still serves as an important piece of qualitative evidence for or against models.

[13] In the domain of monetary losses, the magnitude effect may attenuate (Loewenstein & Prelec, 1992; Scholten & Read, 2013; Scholten et al., 2014), or even reverse (Hardisty et al., 2013a).

2.3.3 Delay effect

The discount rate declines with the delay to the outcome (Thaler, 1981). For example, someone indifferent between $200 in one year and $100 today (an implied interest rate of 100% per year) might prefer $400 in two years to $100 today (less than 100% per year). The delay effect has been one of the most prominent pieces of evidence for the family of hyperbolic discounting models. It must be recognized, however, that the delay *to* the later outcome is confounded with that of the interval *between* the outcomes (Read 2001). All models, but Exponential, can accommodate the delay effect.

2.3.4 Common difference effect and its reversal

The common difference effect is that implied discount rates decrease as the delay to the onset of an interval increases (Loewenstein & Prelec, 1992). For example, someone indifferent between $200 in one year and $100 today might prefer $200 in two years to $100 in one year. The common difference effect has proven less robust than the above phenomena: It is seen in some studies (Green et al., 1994; 2005; Holt et al., 2008; Keren & Roelofsma, 1995; Kirby & Herrnstein, 1995; Read et al., 2005; Scholten & Read, 2006), but not reliably in others (Ahlbrecht & Weber, 1997; Baron, 2000; Holcomb & Nelson, 1992; Read, 2001; Read et al., 2005; Read & Roelofsma, 2003), and, occasionally a *reverse* common difference effect is observed (Attema et al., 2010). Thus, for example, someone indifferent between $200 in two years to $100 in one year might prefer $200 in one year to $100 today. All models, but Exponential and Quasi-hyperbolic, can fully accommodate the common difference effect. Exponential cannot accommodate this effect and Quasi-hyperbolic is just able to accommodate a special case of the common difference effect, the present bias (or immediacy effect), in which the earlier interval has no front-end delay. Only the constant-sensitivity discounting model alone can accommodate the reversal of the common difference effect.

2.3.5 Nonadditivity of intervals

Nonadditivity of intervals can manifest itself in two ways. *Subadditivity* is that implied discount rates are lower over an undivided interval than over its subintervals (Kinari et al., 2009; McAlvanah, 2010; Read, 2001; Read & Roelofsma, 2003, Scholten & Read, 2006, Zauberman et al., 2009). *Superadditivity* is the reverse pattern, in which implied discount rates are higher over an undivided interval than over its subintervals (Scholten & Read, 2006; 2010; Scholten et al., 2014). Consider someone

who is indifferent between (x_S, t_S) and (x_M, t_M), indifferent between (x_M, t_M) and (x_L, t_L), and indifferent between (x_S, t_S) and (y, t_L), where $0 < x_S < x_M < x_L$, and $0 \leq t_S < t_M < t_L$. Subadditivity is that $x_M < y < x_L$, whereas superadditivity is that $y > x_L$, meaning that time has less or more impact, respectively, over the undivided interval than over its subintervals.[14] The hybrid discounting model (Interval) and the full tradeoff model (TM) can accommodate nonadditivity (see Scholten et al., 2014).

2.3.6 Relative nonadditivity

Let $C = \{SS, MM, LL\}$ be a set of single dated outcomes as defined previously, and let $c = \{ss, mm, ll\}$ be another set of single dated outcomes obtained from C by reducing all three outcomes by a common factor. Relative nonadditivity is that subadditivity is more common in c than in C, but superadditivity is more common in C than in c (Scholten & Read, 2010; Scholten et al., 2014). Scholten et al. (2014; pp.420) have shown that relative nonadditivity emerges as a regular pattern of product-rule violations from the combination of TM and a ratio interpretation of Luce (1959) choice axiom (referred to as the Luce specification): In c, the odds of choosing *ll* rather than *ss* are higher than the joint odds of choosing *mm* rather than *ss*, and *ll* rather than *mm* (subadditivity), but, in C, the odds of choosing *LL* rather than *SS* are lower than the joint odds of choosing *MM* rather than *SS*, and *LL* rather than *MM* (superadditivity). We further argue that this feature is applicable to all other attribute-based models involved in this model contest, expect for PD (see Appendix 2D for the proof).

2.3.7 Interim summary

The above review is a sample of behavioural regularities relating to SS-LL choice between monetary gains. We restrict our focus to non-framing behavioural regularities, and do not consider phenomena related to how options are described. Framing-related phenomena in the SS-LL paradigm include the delay-speedup asymmetry (Loewenstein, 1988; Scholten & Read, 2013; Weber et al., 2007), the date/delay effect (LeBoeuf, 2006; Read et al., 2005a), outcome framing effects (Read et al., 2005b; Read et al., 2013) and the (asymmetric) hidden-zero effect (Magen et al., 2008; Read et al., in press; Wu and He 2012). See *framing effects* in Section 1.2.1 in Chapter 1 (pp. 17) for a more detailed review.

[14] See Dai (2017; in press) and Scholten (in press) for recent debates on these effects at both individual and aggregate levels.

2.4 Incomparability of Past Results

There have been many published comparisons between intertemporal choice models. Early work primarily focused on alternative-based models (e.g., Abdellaoui et al., 2010; Cairns & van der Pol, 1997; 2000; Cavagnaro et al., 2016; Kirby & Maraković, 1995; Madden et al., 1999; McKerchar et al., 2009; Peters et al., 2012; Pine et al., 2009; Takahashi et al., 2008; Yi et al., 2009), and a growing number of studies included attribute-based models in the model comparisons (e.g., Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014; Ericson et al., 2015; Scholten & Read 2013; Scholten et al., 2014; Stevens, 2016).

The results from previous model selection studies are mixed and do not agree with each other. Among alternative-based models, many studies suggested that $\text{Hyperbolic}_{\text{H}}$, the one-parameter hyperbolic model, outperformed the exponential discounting model (e.g., Abdellaoui et al., 2010; Kirby & Maraković, 1995; Madden et al., 1999; McKerchar et al., 2009). Some others found that $\text{Hyperbolic}_{\text{LP}}$ outperformed the exponential discounting model and $\text{Hyperbolic}_{\text{H}}$, the one-parameter hyperbolic model (Cairns & van der Pol, 1997; 2000; Mcerchar et al., 2009). Peters et al. (2012) found that the constant-sensitivity discounting model provided the best fit to their experimental data. However, Cavagnaro et al. (2016) suggested that none of the alternative-based models they considered provided satisfactory fits to all behavioural patterns exposed in their results. When attribute-based models were involved in the model comparison, attribute-based models usually outperformed alternative-based or hybrid models. However, different studies included different attribute-based models, thus there is a lack of coherent message from those studies.

We suggest several reasons for the mixed and incoherent evidence from existing model comparisons. The first reason is *model selectivity*: Most studies include a small subset of models, and different studies involve different subsets. Transitivity is not guaranteed, because different models might best fit different data. For example, if one study suggests that Model A outperforms Model B, and another study suggests that Model B outperforms Model C based on a different data set, it is not guaranteed that Model C will outperform Model A in the latter study.

The second reason is *stochastic-specification selectivity*. To test a choice model, it is necessary to specify not just the model, but also how the model converts its outputs into choice probabilities (Table 2.4). Different model tests use different specifications. For example, Scholten et al. (2014) used the Luce specification;

Ericson et al. (2015) used the Logistic specification; and Cheng and González-Vallejo (2016) used different specifications for different models. It is well understood that stochastic specification can make a big difference to the relative accuracy of decision or categorisation models (e.g., Blavatskyy & Pogrebna, 2010 Wills, Reimers, Stewart, Suret, & McLaren, 2000), but is largely overlooked in the literature on intertemporal choice (but see Dai & Busemeyer, 2014, who compared static stochastic specifications, the Logistic and the Fechnerian specifications, with dynamic drift-diffusion models).

Third, *stimulus diversity*: Not only do the models and specifications differ from study to study, so do the stimuli. Outcomes vary from cents to hundreds of thousands of dollars, and delays from days to decades. Also, stimulus designs vary across studies, partly depending on the preference patterns that the researcher plans to expose (e.g., how implied discount rates vary as a function of outcome sign and magnitude, or as a function of delay and interval length).

2.5 Methods

Given the issues with the comparability of past results, the study in this chapter resolves them by applying a large set of fifteen models (in Table 2.2) with three stochastic specifications (in Table 2.4) to 256 data sets obtained from published papers and unpublished projects(see Appendix 2B for the summary of the data sets).

2.5.1 Data sets

Intertemporal choice data were collected from published papers and unpublished projects Specifically, we selected only data of binary choices between single dated monetary outcomes (i.e., SS vs. LL).[15] This ruled out other preference elicitation methods (e.g., matching, pricing, rating, and allocation), non-monetary outcomes (e.g., health and food), and data of choices between outcomes sequences.

We requested data in various ways, although it was not meant to be a meta-analytic literature search. First, a request email was sent out on December 3, 2013 to the Society of Judgment and Decision Making mailing list, asking for aggregate data or raw data from studies that conformed to our selection criteria. Second, one of the authors requested data from intertemporal choice researchers whom he had connections with. Third, we requested data from authors of studies that have cited Kirby et al. (1999), which is one of the most widely cited empirical studies employing

[15] In very few studies, participants were also allowed to be indifferent between the two options.

the SS-LL paradigm, at least once in Google Scholar. Note that not only did we request the data from the specific studies citing Kirby et al. (1999), but also other studies that met our selection criteria from the same authors. Fourth, we downloaded data sets available online from Judgment and Decision Making and American Economic Review with the approval from the author(s) of the papers. Lastly, we searched CNKI (China National Knowledge Infrastructure), a comprehensive Chinese scholarly database in February 2014 and sent request emails to the authors of the identified papers.

We obtained data from 112 published papers and unpublished projects. From each paper or project, separate data sets were created when there were different experiments, when there were different experimental conditions, or when the respondents were residents in different countries. When experimental conditions differed only in outcomes and/or delays used to expose certain phenomena (e.g., the common difference effect), and no other manipulations took place, they were combined in one data set. For example, to test the immediacy effect, Keren and Roelofsma (1995) asked one group of participants to choose between $150 today and $200 in 9 months and another group to choose between $150 in 10 years and $200 in 10 years and 9 months. Although the two questions were answered by different participants, they were jointly used to expose a specific frame-free choice pattern of our interest in this study. Thus, the two items were combined into the same data set. When participants were allowed to be indifferent between the available options, "indifferent" responses were removed, and the sample size for an item was reduced accordingly. Data sets with only one item were removed from the analysis; since models are fit to each data set individually, a single-item data set is uninformative in a model contest. Stimuli in the qualifying data sets were then screened according to the following criteria:

(1) The LL outcome (x_L) was larger than the SS outcome (x_S);

(2) The LL delay (t_L) was longer than the SS delay (t_S);

(3) Both outcomes were to be received with certainty.

Using all these criteria, we were left with 256 data sets from 97 papers or projects (see Appendix 2B for details). Each line in a data set contained the amounts and delays, the frequency with which participants in the data set chose LL, and the number of participants for the item in the data set.

2.5.2 Models

As mentioned earlier, the fifteen intertemporal choice models involved in the contest included eight alternative-based models, six attribute-based models and one hybrid model. The alternative-based models were the exponential discounting model, five variants of the hyperbolic discounting model (Quasi-hyperbolic, $\text{Hyperbolic}_{\text{H}}$, $\text{Hyperbolic}_{\text{M}}$, $\text{Hyperbolic}_{\text{LP}}$ and $\text{Hyperbolic}_{\text{SRS}}$), the constant-sensitivity discounting model, and the double-exponential discounting model. Of the eight alternative-based models, seven were outcome-discounting models and only $\text{Hyperbolic}_{\text{SRS}}$ was a value-discounting model. The six attribute-based models were the proportional-difference model (PD; Cheng & González-Vallejo, 2016), the intertemporal-choice-heuristics model (ITCH; Ericson et al., 2015), the difference-ratio-interest-finance-time model (DRIFT; Read et al., 2013), and three variants of the tradeoff model (BTM_{SR}, BTM_{DB} and TM). The hybrid model had only one member, the discounting by intervals model (Interval).

2.5.3 Stochastic specifications

Three different stochastic specifications were used to model choices between SS and LL. The Luce specification views the strength of preference for LL over SS as a function of the ratio between the value (in alternative-based models) or advantage (in attribute-based models) of LL and the value or advantage of SS. The Logistic specification is transformation of the Luce with the value or advantage of LL and SS exponentiated respectively. The third is the Fechnerian specification, which views the strength of preference for LL over SS as a function of the difference between the value or advantage of LL and the value or advantage of SS, with the difference being evaluated under a standard cumulative normal distribution Φ. The Logistic and Fechnerian specifications, in the functional forms, specify the strength of preference with the difference between the values or advantages of the two options. Thus, they are both termed *difference* specifications.

Table 2.4 List of stochastic specifications.

Name	Stochastic specification
Luce	$\hat{p} = \frac{U_{RHS}^{1/\varepsilon}}{U_{RHS}^{1/\varepsilon} + U_{LHS}^{1/\varepsilon}} = \frac{1}{1 + (U_{LHS} / U_{RHS})^{1/\varepsilon}}$

Logistic	$\hat{p} = \frac{e^{\frac{1}{\varepsilon}U_{RHS}}}{e^{\frac{1}{\varepsilon}U_{RHS}} + e^{\frac{1}{\varepsilon}U_{RHS}}} = L\left(\frac{1}{\varepsilon}\left(U_{RHS} - U_{LHS}\right)\right)$
Fechnerian	$\hat{p} = F\left(\frac{1}{\varepsilon}\left(U_{RHS} - U_{LHS}\right)\right)$

Note. U_{RHS} is the right-hand side, and U_{LHS} the left-hand side, of the equations representing each class of evaluation rules in Table 2.1, and $\hat{p}$ is the predicted probability that LL will be chosen over SS, where $L(\cdot)$ is the standard cumulative Logistic distribution function, $F(\cdot)$ is the standard cumulative normal distribution function, and $\varepsilon > 0$ is the "noisiness" of choice behaviour; as ε approaches 0, choice becomes increasingly determined by the model.

2.5.4 Baseline model

To check whether a data set was informative enough for model selection, a baseline model, in which the predicted probability of choosing LL rather than SS was simply $\hat{p}_s$, i.e., a constant across all items in data set *s*, was involved. If the baseline model turns out to be the winning model in a data set, then that data set was considered uninformative, and removed from the analysis.

2.5.5 Bayesian Information Criterion

Bayesian information criterion (BIC; Schwarz, 1978) was used for model selection. The Bayesian information criterion is a measure of model goodness of fit that penalizes models for the number of parameters it contains. A smaller BIC value indicates better balance between the goodness of fit to data and the complexity of the model. For each combination of a model (*q*), a stochastic specification (*r*), and a data set (*s*), it is given as:

$$\mathrm{BIC}_{qrs} = -2\ln(L_{qrs}) + m_{qr}\ln(n_s),$$

where $\ln(L_{qrs})$ is the maximum log-likelihood for the combination of *q*, *r*, and *s*. The term at the end of the left-had side is a penalty term, where m_{qr} is the number of free parameters in the model specification (*q* and *r*), and n_s is the number of responses across items in data set *s*. The maximum log-likelihood $\ln(L_{qrs})$ is given by:

$$\ln(L_{qrs}) = \sum_{i=1}^{n_s} \ln(B(f_{is}; \hat{p}_{iqrs}; N_{is})),$$

where $B(.)$ is the binomial distribution, f_{is} is the frequency with which participants in data set s choose LL from item i, N_{is} is the number of participants in s choosing either SS or LL from item i, and $\hat{p}_{iqrs}$ is the probability of participants in data set s choosing LL from item i as predicted by the model specifications (q and r). Maximum log-likelihood values were obtained using the Nelder-Mead Simplex Method in Matlab (see Appendix 2A for more details).

2.6 Results

A screening process with the baseline model was applied to select only the data sets that were informative enough for model selection. With the eligible data sets, we will successively discuss the best-performing stochastic specification for each model, model performance when models were combined with the best-performing stochastic specification, and the adverse consequences of a monolithic stochastic specification for all models.

2.6.1 Data screening

For each data set, 46 BIC values were estimated (15 models × 3 stochastic specifications plus 1 baseline model). The baseline model gave the lowest BIC value for 31 data sets, which were removed from the analysis, thus leaving us with 225 eligible data sets.

The stimuli in these eligible data sets differed in various aspects. In the following, we highlight three important and manageable aspects. First, out of the 225 data sets, 144 (64.0%) involved only choices between an immediately available option and a delayed option (intervals without upfront delays). Another 39 data sets (17.3%) involved only choices between two delayed options (intervals with upfront delays). Only 42 (18.7%) involved both choices between immediate-delayed options and choices between delayed-delayed options.

Second, the interest rates implied by the stimuli were much higher than prevailing market rates. Interest rates implied by 16,003 stimuli in 225 data sets, i.e., $r = (x_L / x_S)^{1/(t_L - t_S)} - 1$, where t was measured in days[16], showed a distribution covering a range of very high interest rates (Figure 2.1a). While a daily interest rate of 0.01% is equivalent to 3.7% per year, which is similar to (or still higher than) the prevailing

[16] The implied daily interest rate, rather than the implied annual interest rate, is presented and plotted in the figure because the implied annual interest rate is much more dispersed and skewed (even in a log-scale) and thus less friendly to graphical exposition. The transformation process of some items even exceeded the limit of the double-precision numeric format.

market savings rates, 97% of the implied daily interest rates were higher than 0.01% per day, and 48% of them were higher than 1% per day, which translates to more than 3,678% per year. This observation suggested that, on one hand, researchers generally had the consensus that participants were generally excessively impatient compared with the prevailing interest rates in the market, based on either their intuition or their experience with human participants. On the other hand, because of the context effects such as the background contrast effect (Simonson & Tversky, 1992; Stewart et al., 2015), such stimulus designs could also contribute to the often-observed excessive discounting in the literature (Frederick et al., 2002).

Third, the data sets differed substantially in both the number of stimuli and the average number of participants per item. The number of stimuli in a data set could be as small as 2 and as large as 720 (Figure 2.1b). A small number of stimuli means a limited amount of information in the data, which should benefit models with few parameters, such as the one-parameter models Exponential, HyperbolicH, and PD, due to their parsimony. The average number of participants per item means the weighted sample size of participants in a data set. For example, if, in a data set with 20 items, 17 items had 100 responses each and the remaining three items had 95 responses each, the average number of participants per item was (100 × 17 + 95 × 3) / 20 = 99.25. Across data sets, the average number of participants per item can be as small as 1.5 per item and as large as 3,500 per item (Figure 2.1b).

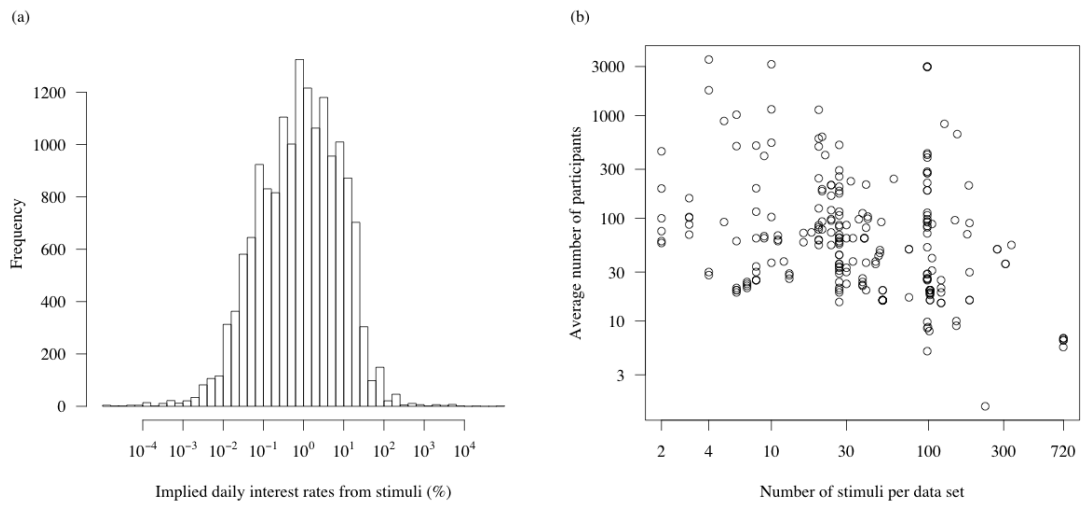

Figure 2.1. Details of the data sets. (a) Histogram showing implied daily interest rates on a log-scale. (b) Scatterplot showing the distribution of the number of items

(x-axis) and the average number of participants (y-axis) in the data sets, both on log-scales.

2.6.2 Stochastic specification influencing model performance

We began the formal analysis by showing the influence of the stochastic specification on model performance. Figure 2.2 provides the comparisons between stochastic specifications. For each model q and stochastic specification r, we computed an Aggregate BIC value (ABIC), meaning BIC values summed across all data sets:

$$\mathrm{ABIC}_{qr} = \sum_{s=1}^{S} \mathrm{BIC}_{qrs},$$

where S = 225 (number of data sets).

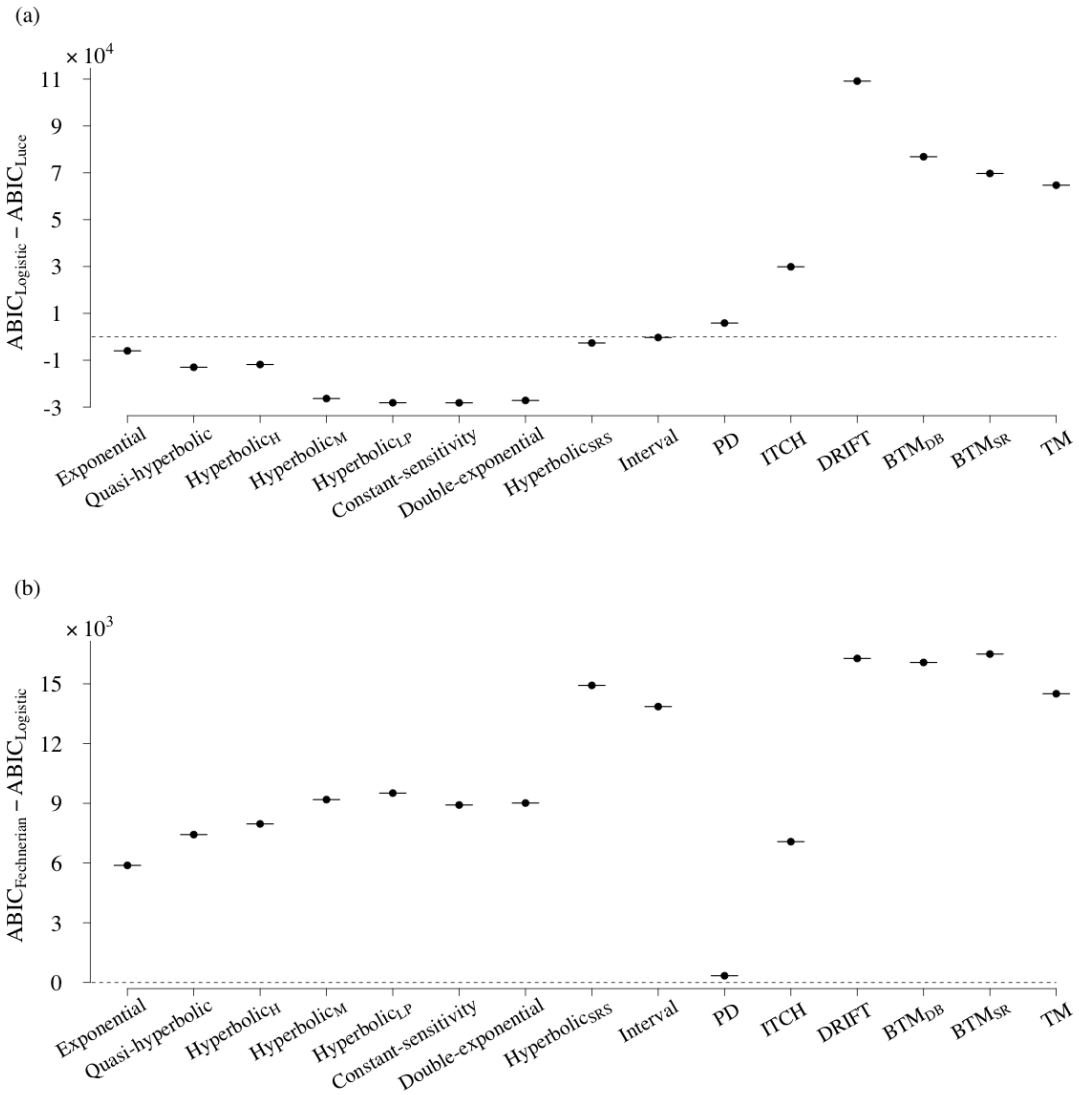

Figure 2.2. Differences in Aggregate BIC values between stochastic specifications. (a) Differences between the Luce specification and the Logistic specification showing that the Luce specification was better for attribute-based models but the Logistic specification was better for alternative-based models and the hybrid model. (b) Differences between the Logistic specification and the Fechnerian specification showing that the former was better than the latter for all models.

Figure 2.2a compares, for each model *q*, the Luce specification with the Logistic specification. Among the alternative-based models, all outcome-discounting models performed better with the Logistic specification than with the Luce specification. The value-discounting models, one of which is an alternative-based model ($\text{Hyperbolic}_{\text{SRS}}$) and the other a hybrid model (Interval), performed about equally well with the two stochastic specifications. All attribute-based models performed better with the Luce than with the Logistic specification. Figure 2.2b compares, for each model *q*, the Logistic specification with the Fechnerian specification. All models performed better with the Logistic than with the Fechnerian specification, although the difference for PD was small in comparisons with the differences for other models.

It was clear that the stochastic specification made a big difference, and that we could not rely on a single specification to models when comparing different models. Thus, as the core interest of this study, the comparison of different intertemporal choice models below allowed the models to be flexibly combined with the stochastic specification that maximised their performance as follows.

2.6.3 Model performance with the best-performing stochastic specification

Given that different specifications produced different ranking of models, the choice of stochastic specification became an issue. To be maximally generous, we allowed each combination of a model and a data set to boost its performance by matching it with the best-performing stochastic specification. For the ease of exposition, we defined the Best BIC, or BBIC, of model *q* for data set *s* as the lowest value of the BIC values produced by the three stochastic specifications:

$$\text{BBIC}_{qs} = \min_{r=1}^{R} \text{BIC}_{qrs}$$

where R = 3 (number of stochastic specifications). The Aggregate BBIC value (ABBIC) of model q was then obtained by summing the BBIC values across all datasets:

$$\mathrm{ABBIC}_q = \sum_{s=1}^{S} \mathrm{BBIC}_{qs} ,$$

where S = 225 (number of data sets). Figure 2.3 shows these Aggregate BBIC values. The three tradeoff models (TM, $\mathrm{BTM}_{\mathrm{DB}}$, and $\mathrm{BTM}_{\mathrm{SR}}$ respectively) came out best from the contest. The other two attribute-based models (DRIFT and ITCH) came out next, followed by the value-discounting models (Interval and $\mathrm{Hyperbolic}_{\mathrm{SRS}}$), in turn followed by the outcome-discounting models. The one-parameter proportional-difference model (PD) came out worst.

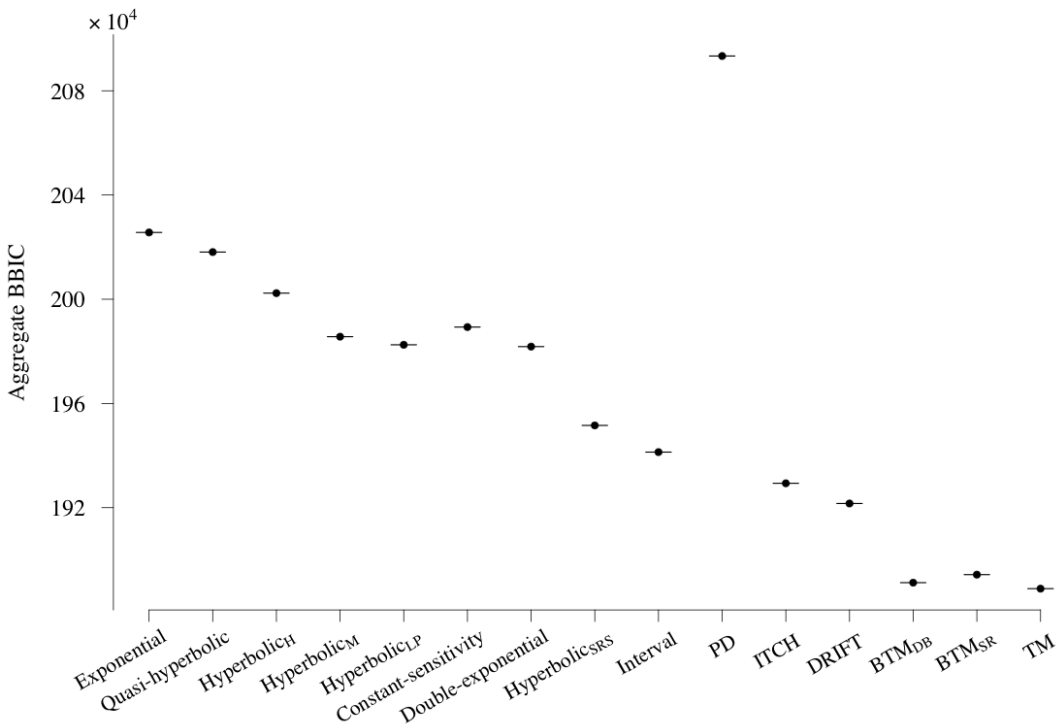

Figure 2.3. Aggregate BBIC value for each model q (ABBIC_q).

One possibility for the victory of the attribute-based models could be that they were favoured by only a small proportion of data sets, but the strength of evidence from them were extremely strong. Thus we continued to test whether or not a majority of data sets favoured the attribute-based models. Table 2.5 presents the percentage of all 225 data sets that each combination of model and stochastic specification provided

the lowest BBIC value for. To be clear, the *lowest BBIC* value is the lowest of the BBIC values across the fifteen models, or equivalently the lowest BIC value across the 45 combinations of model and stochastic specification, for a data set. A breakdown according to the model suggested that attribute-based models achieved the lowest BBIC value for 76.6% of the data sets. Among the discounting models, whether alternative-based or hybrid, the one-parameter specifications achieved the lowest BBIC value for 20.3% of the data sets, with $\text{Hyperbolic}_{\text{H}}$ coming out first (13.4%), and Exponential second (6.8%). Combined, all remaining discounting models achieved the lowest BBIC value for only 3.1% of the data sets.

Table 2.5 Percentage of data sets identifying each combination of model and stochastic specification as producing the lowest BBIC value.

Category	**Model**	**Stochastic specification**			**Total**
		Luce	Logistic	Fechnerian	
Alternative-based	Exponential	4.3%	0.5%	2.1%	**22.5%**
	Quasi-hyperbolic	0.1%	0.0%	0.2%	
	$\text{Hyperbolic}_{\text{H}}$	7.1%	1.7%	4.6%	
	$\text{Hyperbolic}_{\text{M}}$	0.3%	0.0%	0.2%	
	$\text{Hyperbolic}_{\text{LP}}$	0.3%	0.4%	0.0%	
	Constant sensitivity	0.3%	0.0%	0.0%	
	Double exponential	0.0%	0.0%	0.0%	
	$\text{Hyperbolic}_{\text{SRS}}$	0.4%	0.0%	0.0%	
Hybrid	Interval	0.0%	0.0%	0.9%	**0.9%**
Attribute-based	PD	5.1%	4.2%	3.3%	**76.6%**
	DRIFT	11.1%	0.9%	0.4%	
	ITCH	7.1%	1.3%	0.4%	
	BTM_{DB}	20.9%	2.7%	0.0%	
	BTM_{SR}	10.7%	1.3%	1.8%	
	TM	4.0%	0.9%	0.4%	
Total		**71.5%**	**14.0%**	**14.5%**	**100%**

Note. When more than one combination (*m* in number) has the lowest BBIC value, the count is evenly split (1 / *m*) among those combinations.

There are two apparent contradictions in the results presented above. First, the full tradeoff model (TM) outperformed all other attribute-based models in terms of

Aggregate BBIC values (Figure 2.3), but was outperformed by all those models in terms of the percentage of data sets that identified them as producing the lowest BBIC value (Table 2.5). Second, Exponential and $\text{Hyperbolic}_{\text{H}}$ seemed to beat all other discounting models, whether alternative-based or hybrid, regarding the percentage of data sets that they provided the lowest BBIC values (Table 2.5), but were beaten by all other discounting models (except for the quasi-hyperbolic discounting model) on the Aggregate BBIC values (Figure 2.3). To foster a better understanding of the contradictions, we carried out a more in-depth analysis in two separate waves for the attribute-based models and the (alternative-based and hybrid) discounting models respectively.

Pairwise Comparisons between attribute-based models. To clarify this contradiction among attribute-based models, we disaggregated the data sets, and compared, for each data set *s*, the BBIC value achieved by the full tradeoff model (TM) with the BBIC values achieved by other attribute-based models. Figure 2.4a compared TM with basic tradeoff model by Scholten and Read (2013; BTM_{SR}). A majority of the differences between the BBIC values (89.3%) fell above 6 (in favour of BTM_{SR}) or below −6 (in favour of TM), which, by a conventional standard (Kass & Raftery, 1995; Raftery, 1995), constituted "strong" support or beyond for the winning model and there were many more occasions on which BTM_{SR} received the support (85.1%) than occasions on which TM did (14.9%). However, if we look at the differences falling above 100 or below −100, which by an informal standard constitute "extremely strong" support for the winning model, there were eight occasions on which TM received the support (88.9%) but only one occasion on which BTM_{SR} did (11.1%). Additionally, while the strongest evidence for BTM_{SR} was a BBIC difference of −119, the strongest evidence for TM was a BBIC difference as high as 4,731. Therefore, we were confident to tell that BTM_{SR} was favoured by the number of matches won, whereas TM was favoured by its extremely high scores in a relatively small number of matches won. Figures 2.4b-d tell similar stories about the comparisons between TM and BTM_{DB}, DRIFT and ITCH respectively. Figures 2.4e, however, suggests that TM outperformed PD with respect to both the number of matches won and the strength of evidence in the matches won.

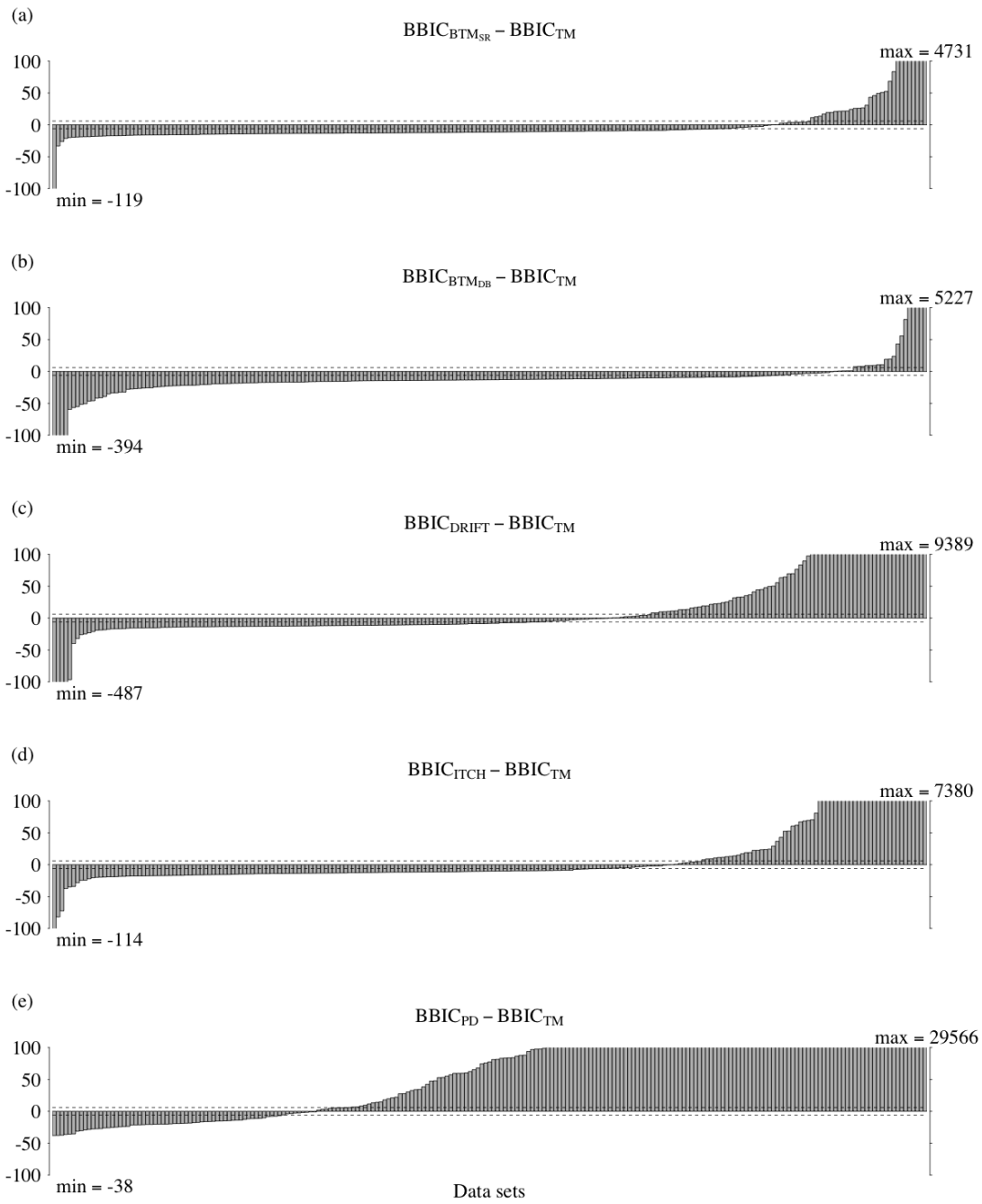

Figure 2.4. Pairwise differences in BBIC values (BBIC$_{qs}$) between the full tradeoff model (TM) and other attribute-based models. The 225 data sets are ordered from the most negative difference (favouring another attribute-based model other than TM) to the most positive difference (favouring TM). Dashed lines are positioned at BBIC differences of 6 or -6.

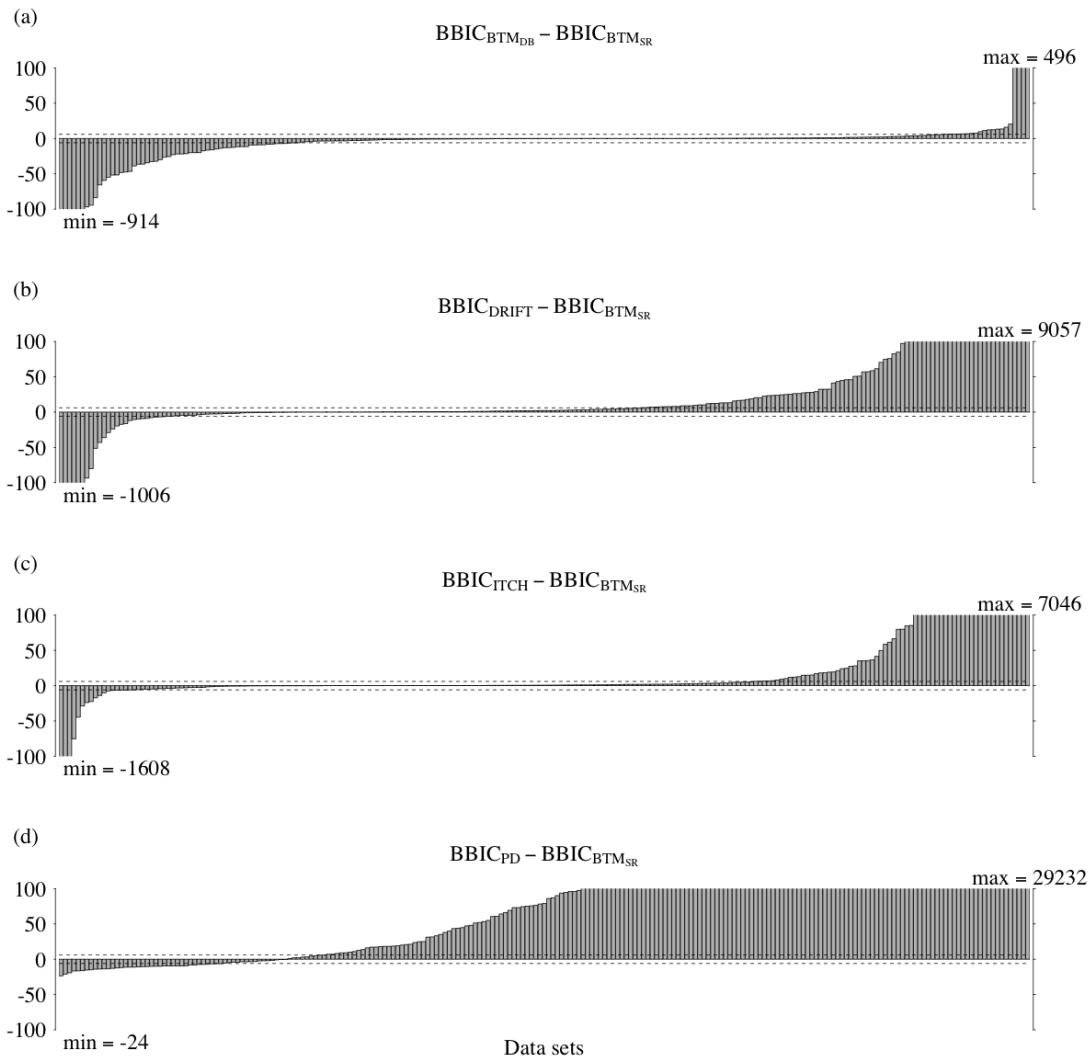

Figure 2.5. Pairwise differences in BBIC values ($BBIC_{qs}$) between BTM_{SR} and other attribute-based models. The 225 data sets are ordered from the most negative difference (favouring another attribute-based model) to the most positive difference (favouring BTM_{SR}). Dashed lines are positioned at BBIC differences of 6 or -6.

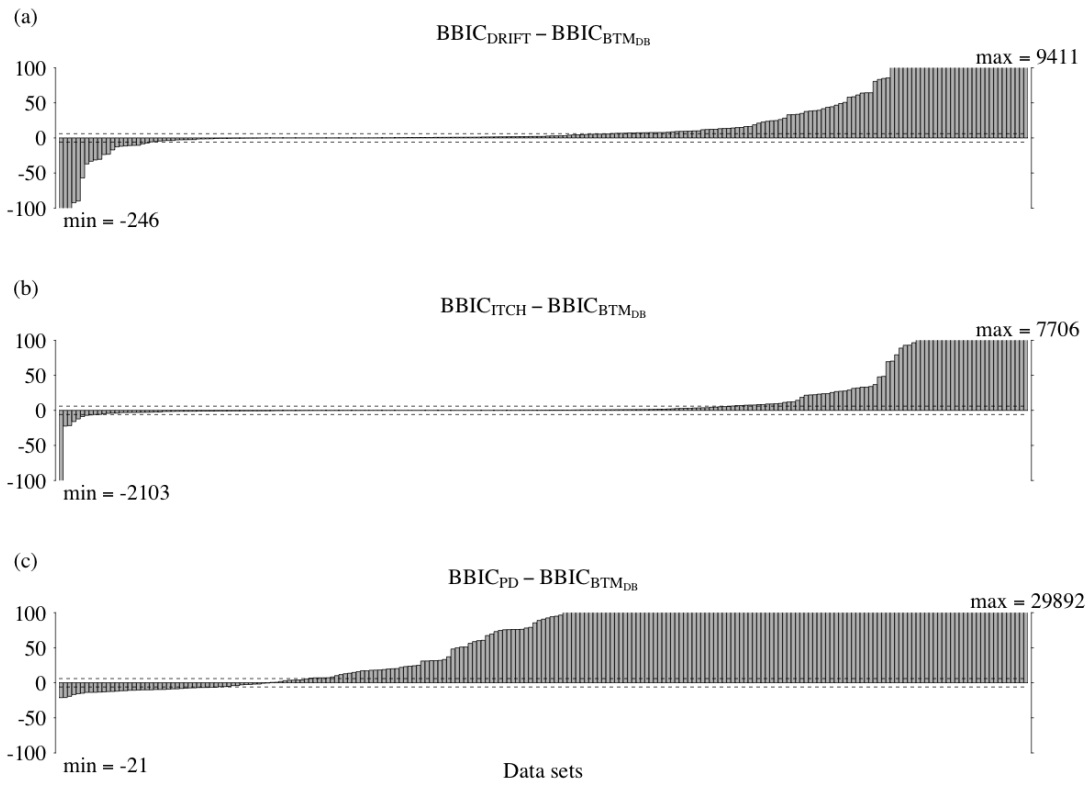

Figure 2.6. Pairwise differences in BBIC values ($BBIC_{qs}$) between BTM_{DB} and attribute-based models. The 225 data sets are ordered from the most negative difference (favouring another attribute-based model) to the most positive difference (favouring BTM_{DB}). Dashed lines are positioned at BBIC differences of 6 or -6.

Further pairwise comparisons between the basic tradeoff models (BTM_{SR} and BTM_{DB}) and other attribute-based models suggested these two models outperformed other attribute-based models (see Figure 2.5 and Figure 2.6). Finally, a pairwise comparison between the two basic tradeoff models (BTM_{SR} and BTM_{DB}) suggested BTM_{DB} had an edge over BTM_{SR}. Particularly, they performed almost equally well (|BBIC difference| < 2) for 50.7% of the data sets and another 16.0% provided only modest support to one or the other (2 ≤ |BBIC difference| < 6), but, for the remaining data sets, 73.3% of the remaining datasets (55 of 75) provided strong support to BTM_{DB} and only 26.7% (20 of 75) provide strong support to BTM_{SR} (see Figure 2.5a).

Comparisons between discounting models. For the contradiction among the discounting models, certainly, the Aggregate-BBIC criterion should be more reliable than the Percentage-of-the-Lowest-BBIC criterion because the latter was silent on the majority of data sets that lent their support exclusively to the attribute-based models.

Indeed, the former was also consistent with the findings from previous intertemporal choice model contests that Exponential and Hyperbolic$_H$ were typically outperformed by other discounting models (e.g., Abdellaoui et al., 2010; Cavagnaro et al., 2016; Peters et al., 2012; Scholten et al., 2014; Takahashi et al., 2008).

Exponential and Hyperbolic$_H$ were the only one-parameter discounting models in the contest and were special cases for many other discounting models. Thus they could win out only because of their parsimony, especially when data sets had a small number of stimuli or showed a narrow array of preference patterns. The boxplots in Figure 2.7 confirms this conjecture by showing that the subsets of data lending support to the one-parameter discounting models (i.e., Exponential and Hyperbolic$_H$), as well as the one-parameter PD model, generally had less items than other data sets. When data sets showed a broader array of patterns, other discounting models would probably fare better than the one-parameter discounting models (as indicated in Figure 2.3), but, in the majority of such cases, attribute-based models fared better still, with only a total of 3.1% of the lowest BBIC values going to the multi-parameter discounting models.

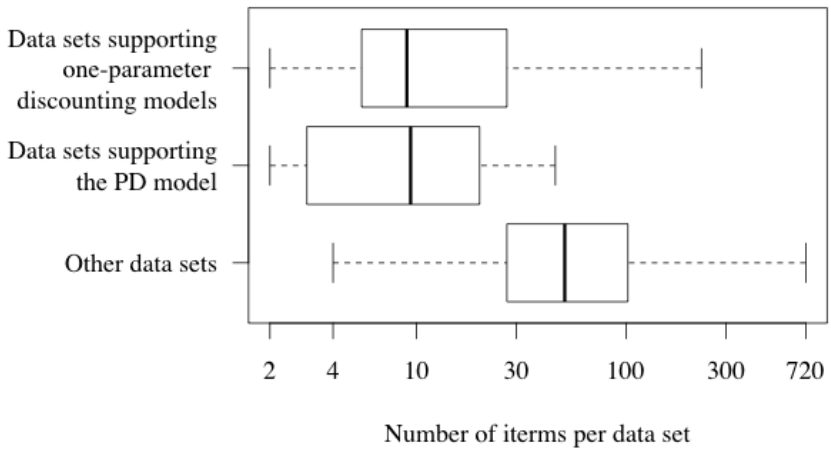

Figure 2.7. Boxplots of the number of items per data set (*x*-axis is on a log scale). All eligible data sets are divided into three subsets: supporting one-parameter discounting models (n = 51), supporting the PD model (n = 34), and others (n = 151). The numbers do not sum up to 225 because 11 data sets appear in the first two categories both.

2.6.4 Model performance with monolithic stochastic specifications

We have so far seen that the attribute-based models, especially the three variants of the tradeoff model, emerged as the winning models (Figure 2.3 and Table

2.5) and that, from comparisons of Aggregate BIC values between stochastic specifications, the Luce specification emerged as the optimal stochastic specification for attribute-based models, whereas the Logistic specification emerged as the optimal for the alternative-based models and the hybrid model (Figure 2.2). We now evaluate the potential costs of arbitrarily assuming a particular stochastic specification for all models in the contest. Figure 2.8 shows the Aggregate BIC values obtained with the three stochastic specifications respectively. The ranking according to the Aggregate BIC values obtained with the Luce specification was similar to the ranking according to Aggregate BBIC values. However, the Aggregate BIC values with the Logistic specification did not achieve the same resemblance. Instead, with the Logistic specification, the two value-discounting models (i.e., Hyperbolic_{SRS} and Interval) came out best from the contest, overturning the pattern obtained with the Aggregate BBIC values. The Fechnerian specification obtained a similar pattern of ranking to the one by the Logistic specification. We therefore confirmed what Stott (2006) and Blavatskyy and Pogrebna (2010) earlier found outside the domain of intertemporal choice that different stochastic specifications could yield different rankings of models, echoing the vital importance of stochastic specification in choice modelling.

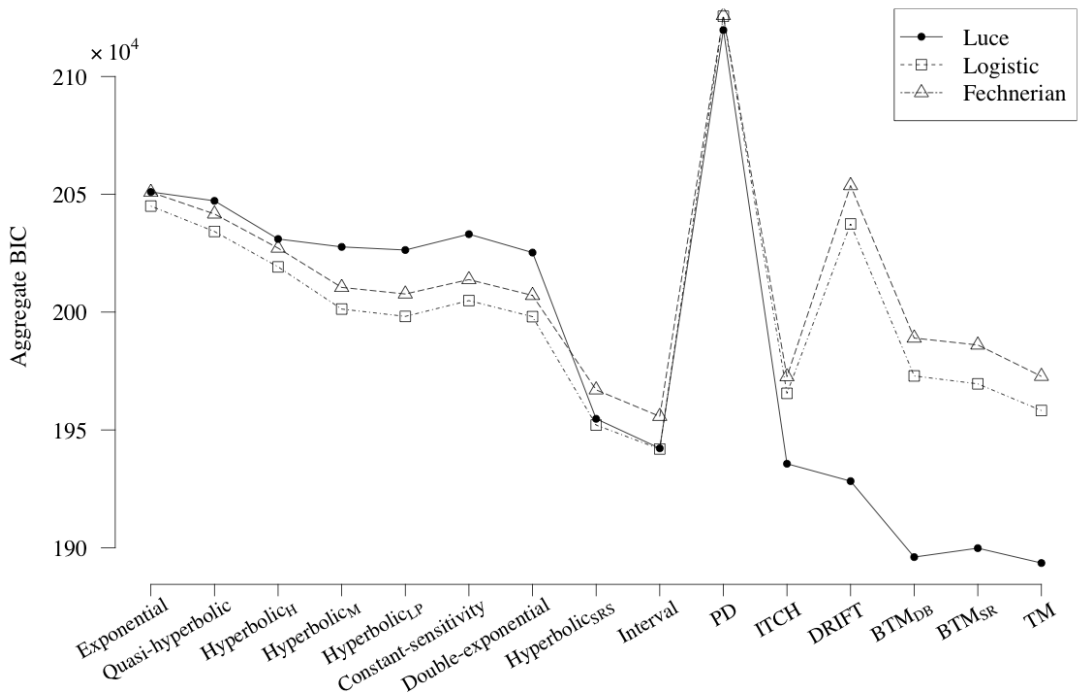

Figure 2.8. Aggregate BIC values with monolithic stochastic specifications.

2.7 Discussion

In this study, the attribute-based models, especially the family of tradeoff models, emerged as the convincing winners from a contest in which fifteen intertemporal choice models, each combined with three stochastic specifications, battled it out over 16,003 items in 225 data sets. In pairwise comparisons between attribute-based models, the full tradeoff model (TM) won fewer matches than most of other attribute-based models, but scored much better in matches won, meaning that, although the TM did not win out very often, if it won out, it won out big. This study also suggested that stochastic specifications played a vitally important role in the performance of intertemporal choice models in the model contest. While Logistic and Fechnerian specifications gave rise to value-discounting models (i.e., $\text{Hyperbolic}_{\text{SRS}}$ and Interval), the Luce specification gave rise to the class of attribute-based models (except for PD), highlighting the importance of stochastic specifications in intertemporal choice modelling.

2.7.1 Attribute-based models as the convincing winner

This study provides extremely strong evidence for the attribute-based models. This evidence also lent support to the attribute-based evaluation rule recently proposed in intertemporal choice (Leland, 2002; Rubinstein, 2003; Scholten & Read, 2010) and, more generally, judgment and decision making (Tversky, 1969; 1972). Further comparisons suggested that attribute-based models outperformed other attribute-based models (see Figure 2.3; this is also confirmed by the pairwise comparisons of BTM_{SR} and BTM_{DB} with other attribute-based models in Figure 2.5 and Figure 2.6).

The proportional difference model (PD; Cheng & González-Vallejo, 2016) is the only exception. The PD model came out poorly in terms of Aggregate Best BIC values, lagging behind all other models, although the percentage of data sets identifying it as the best model was not too bad (note that this percentage score was unreliable in this case as it was silent on the majority of data sets that support other models). These results could be appreciated in the light of two considerations. First, the PD model does not accommodate a basic discounting pattern implied by *positive discounting* that when x_S and t_S are held constant, an increase in $t_L - t_S$ requires an increase in x_L for a person to remain indifferent between SS and LL (Scholten & Read, 2013). Meanwhile, the majority of the data sets in this contest involved only choices between an immediately available option and a delayed option, thus this theoretical

limitation place PD in an obviously disadvantaged position. Second, PD does not accommodate the absolute magnitude effect, which is one of the most robust observations in intertemporal choice studies.[17] With these theoretical limitations, especially the former, it is not surprising that PD perform poorly in the contest.

Despite the overwhelming success of attribute-based models in this model contest, they did not win in a sizeable minority of data sets. Furthermore, these data sets typically involved a few trials each. One possibility is that there is a shift of decision making strategies across trials. In particular, in the starting trials, people may use the alternative-based evaluation rule. When they get more experienced with some trials, they may switch to more heuristic attribute-based evaluation rule. As a result, data sets with more trials are likely to favour attribute-based models and those with less trials are likely to favour alternative-based models. However, the evidence for this proposal is weak in this study, because those data sets with a few trials mostly favoured the simplest alternative-based models primarily because of parsimony.

2.7.2 The importance of stochastic specifications in choice modelling

Like what Stott (2006) observed in risky choice, this study revealed an interaction between model and stochastic specification on model performance: Attribute-based models performed better with the Luce specification, but (alternative-based and hybrid) discounting models performed better with the Logistic specification. We suggested two driving forces that could lead to the interaction. First, the Luce specification capacitates attribute-based models (except for PD) to accommodate the intricate pattern of relative nonadditivity, in terms of the violation to the product rule, while Logistic specification does not. Scholten et al. (2014, pp.420) demonstrate this feature in a basic tradeoff mode (BTM_{SR}). We further argued that this feature was applicable to all the attribute-based models involved in this model contest, expect for PD (see Appendix 2D). This feature may make the Luce specification superior to the Logistic specification for attribute-based models. Second, the Logistic specification capacitates the outcome discounting models, which by themselves do not accommodate the absolute magnitude effect, to produce a *quasi*-absolute magnitude effect when LL is preferred to SS while the Luce specification never does. It is called

[17] The absolute magnitude effect could be accommodated if one sets the threshold ψ at different levels for the different outcome magnitudes (Cheng & González-Vallejo, 2016), as some have done for $Hyperbolic_H$ to accommodate the same effect (e.g. Giordano et al., 2002; Green et al., 1997; Kirby, 1997), but this is not explicitly written in the model and would have led to an endless proliferation of parameter settings within and across data sets.

a quasi-effect because it only predicts a change in the strength of preference, but not a preference reversal. For example, if one prefers \$200 in 2 years to \$100 in 1 year. Under the Logistic specification, that preference can be written as $0 < d(2)200 - d(1)100$, where d (.) is the discount function. It follows that the strength of preference for LL will be augmented if the magnitude of the outcomes is raised by a common ratio k ($k > 1$), i.e., $0 < d(2)200 - d(1)100 < d(2)200k - d(1)100k$. However it is important to note that they could also produce a *reverse-quasi*-absolute magnitude effect when *SS* is preferred to LL, i.e., $d(2)200k - d(1)100k < d(2)200 - d(1)100 < 0$.

We also compared the Logistic specification with the Fechnerian specification, both *difference* specifications. All models performed better with the Logistic specification than with the Fechnerian specification. The cumulative Logistic and the cumulative normal functions are very similar except that the former asymptotes less sharply, or has fatter tails, than the latter (Figure 2.9). Thus, the Logistic specification is more tolerant of choice uncertainty than the Fechnerian specification, and may therefore contribute to better model fits. These findings highlighted the importance of developing alternative stochastic specifications that conveys distribution-free strength of preference in choice modelling (e.g., Cavagnaro, Regenwetter, & Popova, 2016).

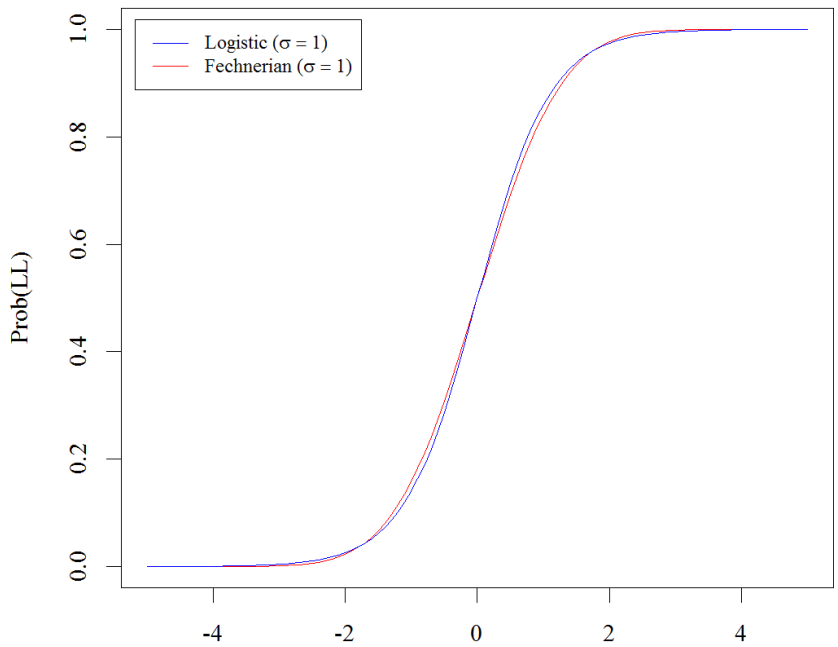

Figure 2.9. The Logistic (standard cumulative logistic distribution) and the Fechnerian (standard cumulative normal distribution) specifications.

2.7.3 Conclusion

Wrapped up, this model contest on a vast amount of secondary data strongly supports attribute-based models, especially the family of the tradeoff model, over alternative-based models of intertemporal choice and the Luce specification over difference specifications as the stochastic specification. Extensions to monetary losses, nonmonetary outcomes, and outcome sequences, whether monetary (Scholten et al., 2016), nonmonetary (Loewenstein & Prelec, 1993), or both (Prelec & Loewenstein, 1998), are recommended, so as to evaluate the robustness of attribute-based intertemporal choice.

CHAPTER 3 ATTENTION AND INTERTEMPORAL CHOICE

Behavioural theories suggest that attention plays an important role in value-based decision making (Bhatia, 2014; Bardalo, Gennaioli, & Shleifer, 2012; Busemeyer & Townsend, 1993; Kőszegi &Szeidl, 2013; Tsetsos, Usher, & Chater, 2010; Tversky, 1972). The more we attend to something the more important it becomes, or vice versa. To quantify the attention effect, some researchers proposed that, based on the drift-diffusion models, attention modulated the speed of value accumulation of options: An option accumulates its value faster when it is focused attention on than when it is not (Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011). Thus, for example, in a binary choice, the more an option is attended to, the more likely it is to be chosen (as long as it has a positive value). Studies on such option-wise attention effects have gained increasing popularity in the literature (Armel, Beaumel, & Rangel, 2008; Franco-Watkins, Mattson and Jackson, 2016; Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011; Shimojo, Simion, Shimojo, & Scheier, 2003; Stewart, Hermens, & Matthews, 2016; Störmer & Alvarez, 2016).

However, the option-wise attention effect is silent on what information people are accessing when focusing attention on an option. Studies have suggested that people could query information from different perspectives, not necessarily in an option-wise way when faced by a decision (Weber, Johnson, Milch, Chang, Brodscholl, & Goldstein, 2007; Tversky 1972). For one, many studies suggested that value-based choice could be made in an attribute-based fashion (Arieli et al., 2011; Leland, 2002; Rubinstein, 2003; Scholten & Read, 2010; Tversky 1972; see Vlaev et al., 2011 for a review). In line with the attribute-based perspective, some studies suggested that attending to (the comparison along) an attribute across options produced an attribute-wise attention effect (Fisher & Rangel, 2014; Experiment 2; Hare, Malmaud, & Rangel, 2011). For example, in a choice between fruit salad and cheesecake, focusing attention on calories will increase the likelihood of choosing the fruit salad against cheesecake while focusing attention on their tastiness will do the reversal. For another, focusing attention on (the valuation) of a single component, or an attribute value in a single option, could produce a component-wise attention effect (Fisher & Rangel,

2014; Experiment 2). For instance, focusing attention on the calories of the cheesecake is different from focusing attention on its tastiness and is also different from focusing attention on the calories of the fruit salad. The option-wise view on attention-driven preference is silent on either attribute-wise or component-wise attention effect.

Based on the existing literature on the relationship between attention and choice, we described an overarching theoretical framework of the attention effects on intertemporal choice and proposed a novel procedure of attention manipulation that allowed us to investigate different ways of attention effect on intertemporal choice.

3.1 Existing Literature

There are two main approaches to studying the attention effects on value-based decision making. First, the attention effect can be measured using eye-tracking, and, second, attention can be manipulated directly.

Eye-tracking studies. In eye-tracking studies, measures of visual attention such as eye movements and gaze duration are directly observed. Most eye-tracking studies focused on the option-wise attention effect (e.g. Fiedler & Glőckner, 2012; Franco-Watkins et al., 2016; Krajbich & Rangel, 2011; Krajbich et al., 2010; Stewart et al., 2016). These studies invariably found an option-wise attention effect on value-based decision making that the more often an option was attended to, the more likely it was to be chosen (when options have positive values). For example, in an eye-tracking study of the attention effect on intertemporal choice between smaller-sooner (SS) and larger-later (LL) options, Franco-Watkins et al. (2016) found that focusing attention on LL increased the likelihood of choosing LL and focusing attention on SS decreased the likelihood of choosing LL.

Fisher and Rangel (2014; Experiment 1) was an exception. They analysed the attention effects on intertemporal choice in a component-wise way: Gaze durations on each of the four components (i.e., SS outcome, SS delay, LL outcome and LL delay) were identified independently, all of which served as the independent variables of a regression with the individual discount rate as the dependent variable. They found that attending to LL outcome decreased the individual discount rate (i.e., increased the likelihood of choosing LL) and attending to SS delay increased the individual discount rate (i.e., decreased the likelihood of choosing LL), while attending to the other two components had no significant behavioural effects. To account for this finding, Fisher and Rangel (2014) proposed a component-wise attention effect: attending to different components in the same option or along the same attribute can contribute

independently to the final choice. However, their conclusion was premature because they did not consider the possibility of the co-existence of multiple ways of the attention effect on value-based decision making, which will be discussed later.

To sum up, eye-tracking generally can be used to test multiple ways of the attention effects on value-based decision making, although this advantage was rarely taken. Despite this advantage, eye-tracking studies can only draw a correlational relationship between attention and choice/preference and thus are silent on whether attention drives or merely reflects preference (Shimojo et al., 2003; Stewart et al., 2016).

Attention-manipulation studies. To draw a causal relationship between attention and value-based decision making, a few studies experimentally manipulated attention and tested the attention effects on decision making. Again, most of them tested the option-wise attention effect. For example, Shimojo et al. (2003) used the *gaze-manipulation* paradigm to study how attention influenced the judgment of facial attractiveness. They presented their participants with a pair of two faces sequentially on a computer screen. The faces were shown alternatively for 300 ms and 900 ms respectively for the same number of repetitions, after which participants decided which one of the faces was more attractive. Option-wise attention was manipulated by varying the duration for which options were directly attended to (i.e., 300 ms vs. 900 ms per repetition), with a longer time inferred as greater attention. Armel et al. (2008) applied the same *gaze-manipulation* paradigm in choices between food items and choices between art posters respectively. A recent study by Störmer and Alvarez (2016) used the psychophysical *attentional-cuing* paradigm (Carrasco, Ling, & Read, 2004) to study the option-wise attention effect on perceived facial attractiveness. Participants were shown two faces simultaneously (for 58 ms) and then decided which one was more attractive. Attention was manipulated by a task-irrelevant attention cue (a black dot for 70 ms) at the location of one of the faces preceding the onset of face presentation. They found that the cued face was more likely to be judged as more attractive.

A few studies experimentally manipulated attribute-wise attention. To elicit attribute-wise attention, Hare et al. (2011) *explicitly asked* participants to attend to the healthiness, the taste or whatever features that automatically came to their minds in different conditions respectively in two-option food choices. This is an attribute-wise attention manipulation because healthiness and tastiness are different attributes or

dimensions in choosing between items of food. Fisher and Rangel (2014, Experiment 2) used the *gaze-manipulation* paradigm to study the attribute-wise attention effect on intertemporal choice. In a standard intertemporal choice task between SS and LL, they presented participants with different attributes (i.e., outcomes and delays) alternatively. Attention was manipulated by presenting one attribute for 2,000 ms, the other for 500 ms alternatively, each repeated twice. They found that longer visual exposure, or greater attention, to the outcomes (i.e., SS and LL outcomes) than to the delays (i.e., SS and LL delays) increased the likelihood of choosing LL and thus decreased impatience. Specifically, when outcomes are exposed longer the proportion of LL choice is 53.6%, opposed to 51.6% when delays longer.

These studies, without an exception, only elicited either the option-wise or the attribute-wise attention effect. Moreover, no previous study elicited the component-wise attention. When the attention effect on value-based decisions operates in different ways, it is of principal interest in this study to synthesize these findings and to provide an integral account of them.

3.2 Framework of the Attention Effects

To help organise the existing literature, we drew up a framework of the attention effects in the context of intertemporal choice. The framework includes option-wise, attribute-wise and component-wise attention effects.

Figure 3.1 provides an overview of the framework. The option-wise effect is that focusing attention on the valuation of an option make the option in focus more likely to be chosen (Figure 3.1a). So focusing attention on SS decreases the chance for LL to be chosen while focusing attention on LL increases the chance that LL be chosen.

The attribute-wise effect, instead, is that focusing attention on the comparison along an attribute makes the attribute overweighed and thus the option favoured by the comparison under attentional focus becomes more likely to be chosen. So focusing attention on outcomes increases the likelihood of choosing LL while focusing attention on delays decreases the likelihood of choosing LL (Figure 3.1b).

The component-wise effect is that focusing attention on the judgment of a component of an option looms larger this specific component. If the component is favourable (i.e., positive in valence), the likelihood that the corresponding option be chosen increases. If the component is aversive (i.e., negative in valence), the likelihood that the option be chosen decreases. Thus, focusing attention on LL outcome [or SS

outcome] decreases [or increases] the likelihood that LL be chosen. By contrast, focusing attention on LL delay [or SS delay] increases [or decreases] the chance for LL to be chosen (Figure 3.1c).

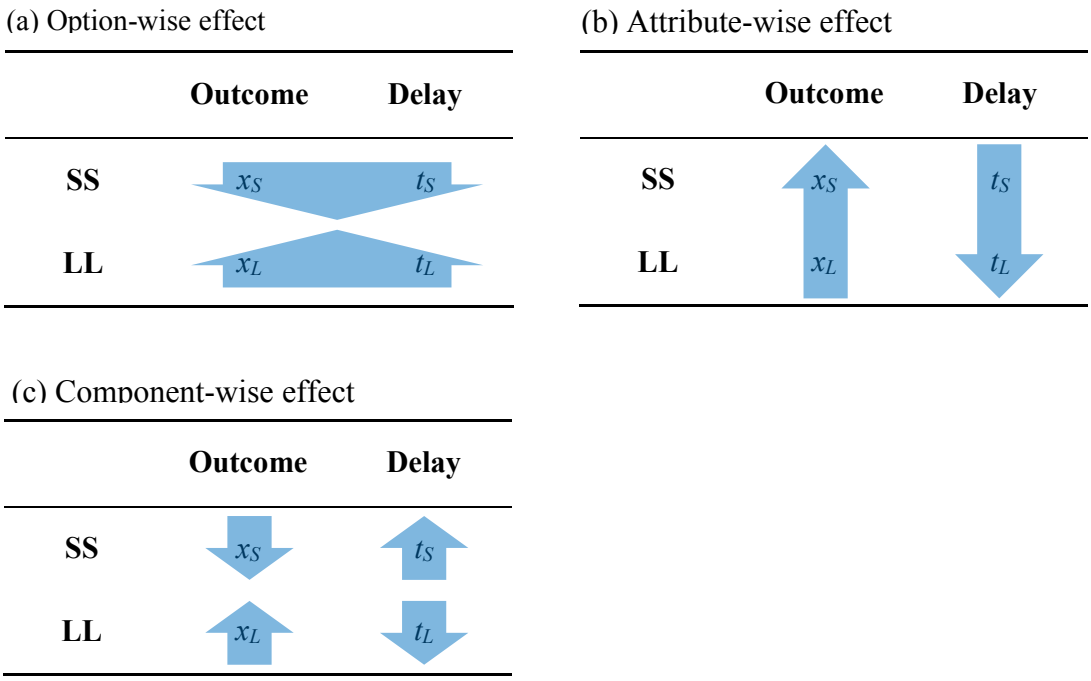

Figure 3.1. Graphical illustrations of different ways of the attention effect on intertemporal choice. The upward arrows mean increased preferences for LL when the corresponding aspects are attended to. The downward arrows mean decreased preference for LL when the corresponding aspects are attended to. (a) Option-wise effect: attention being focused on options. (b) Attribute-wise effect: attention being focused on attributes. (c) Component-wise effect: attention is operated on each component independently from each other.

In addition, different ways of the attention effects (i.e., option-wise, attribute-wise and component-wise) are not exclusive. It is natural that people could allocate attention from different perspectives from time to time when faced with a decision (Busemeyer & Townsend, 1993; Scheibehenne, Rieskamp & Wagenmakers, 2013; Shafir, Simonson, & Tversky, 1993; Tversky, 1972; Weber, Johnson, Milch, Chang, Brodscholl, & Goldstein, 2007). Thus, it is theoretically possible that multiple ways of attention effects could co-exist in the making of value-based decisions. For example, one may first focus attention on the comparison along the outcomes

(attribute-wise) first and then switch her attention to the SS delay alone (component-wise) in an intertemporal choice.

3.3 Attention Manipulation and the Present Study

To study different ways of the attention effects on intertemporal choice, we proposed a new way to manipulate attention in intertemporal choice. Roughly speaking, attention was manipulated by keeping one component varying across items while all other components were kept unchanged (see the experimental procedure of Experiment 1 in Figure 3.2 for an example). Given the long-standing assumption that stimuli that are varied within an otherwise stationary environment draw disproportionate attention (Chun, Golomb, & Turk-Browne, 2011; Franconeri & Simons, 2003; Weber & Johnson, 2009), the varying component should attract more attention than others. For example, in an intertemporal choice between SS and LL, when the LL outcome kept varying across items but other components were kept constant, we expected LL outcome will attract more attention than others.

Manipulating attention in this way had three advantages. First and foremost, this manipulation allowed attention to operate in different ways at the participants' own will. This was a stark contrast to existing paradigms, in which participants were merely to show a pre-determined way of the attention effect on decision making, either option-wise (Armel et al., 2008; Shimojo et al., 2003) or attribute-wise (Fisher & Rangel, 2014; Hare et al., 2011). Second, the procedure opened the door to testing the component-wise attention effect, which had not been achieved by existing paradigms of attention manipulation. Third, the procedure delivered full information of the choice question to participants simultaneously so that the manipulation of attention should not intrude on the natural decision process. As it has been controversial whether people make option-based or attribute-based evaluation in value-based decision making (Leland, 2002; Rubinstein, 2003; Scholten & Read, 2010; Tversky, 1969; 1972), it is essential to display all relevant information to facilitate the natural decision process. By contrast, in the widely-used gaze-manipulation paradigm, only partial information of the choice questions was displayed once at a time, which could intrude the natural decision process.

A drawback of this method was that it could lead to systematically different decision contexts across conditions at the item level, which could produce a background contrast effect (Ebert & Prelec, 2007; Priester, Dholakia, & Fleming, 2004; Simonson & Tversky, 1992). The background contrast effect is a well-

documented violation of sequential independence, which means that the current tradeoff will be influenced by other similar tradeoffs that one has been recently exposed to.[18] This effect could take place when a series of LL options were presented sequentially. While systematically different blocking of items was used to manipulate attention, it could also produce a background contrast effect as a by-product, although the unintended manipulation of background contrast is orthogonal to the manipulation of attention as detailed later. Thus, while testing the attention effect on intertemporal choice, we in the meantime quantified the background contrast and statistically controlled the background contrast effect.

Two experiments were conducted to test the attention effects on intertemporal choice with the new way of attention manipulation. Attention was directed to different components (the outcome and the delay) of LL (Experiment 1) and SS (Experiment 2) respectively.

3.4 Experiment 1: Varying LL Components

In Experiment 1, attention was manipulated by varying either the LL outcome [*outcome-focus* condition] or the LL delay [*delay-focus* condition] across items respectively. As attention was directed to the same option across conditions, the option-wise attention effect was silent on the difference between conditions. The attribute-wise and the component-wise effects both predicted that focusing attention on the LL outcome (in the outcome-focus condition) made people less impatient and thus more likely to choose LL than focusing attention on the LL delay (in the delay-focus condition; see Figure 3.1).

3.4.1 Experiment 1: Methods

Participants. Participants for Experiment 1 were 191 USA residents recruited from Prolific Academic (73 females and 118 males). They were on average aged 31.6 years old. 60.2% of them had a university degree or higher. After completing the experiment, each of them received a flat payment of $1.

Intertemporal choice task. The intertemporal choice task included 81 target items. Across items, the smaller-sooner (SS) option was held constant, which was always receiving $110 in 2 months. The target items were created by crossing nine different LL outcomes with nine different LL delays. The LL outcome could take nine

[18] While referring to similar concepts, Stewart et al. (2003) and Vlaev, et al. (2007) use the term "prospect relativity" and Ebert and Prelec describe this as attention-driven sensitivity to an attribute. We follow Simonson and Tversky (1992) by using the term "background contrast".

values from $120 to $200 by an increment of $10. The LL delay could also take nine different values from 4 months to 20 months by an increment of 2 months. Nine additional screening items with a dominant option, which differed across conditions, were added to the target items to check whether participants were just responding randomly. So, there were totally 90 intertemporal choice items in each condition.

Design. A between-participant design was adopted in Experiment 1. All participants answered the same target intertemporal choice questions, with only the pattern of presentation being different across conditions. In the *outcome-focus* condition, the 81 target items were divided into nine blocks so that only one LL delay, but all LL outcomes, appeared in a block. In each block, an additional screening item was added, whose LL outcome was $100 and LL delay was the same as the LL delay of target items in the same block (e.g., $100 in 6 months), making LL dominated by SS. The screening items were chosen in order not to intrude the one-attribute-varying property within a block. To consolidate the manipulation of attention, the inter-item substitution of LL outcomes was accompanied by a smooth transition animation (1000 ms) with *Odometer*, a jQuery plugin in HTML (see Figure 3.2 for an illustration).[19] The order of blocks and the order of LL outcomes within each block were randomized among participants.

In the *delay-focus* condition, items were also divided into nine blocks of nine items each. In each block, all LL delays, but only one LL outcome, appeared. An additional screening item was added to each block. The LL outcome of the screening item was the same as the corresponding LL outcome in the block and the LL delay is 2 months (e.g., $150 in 2 months), making LL dominate SS. Within each block, the LL delay varied with the same inter-item transition animation (1000 ms), while others were kept constant across items. The order of blocks and the order of items in each block were randomized across participants.

[19] The link to Experiment 1 is: http://gilum.lnx.warwick.ac.uk/wbs/experiment1/

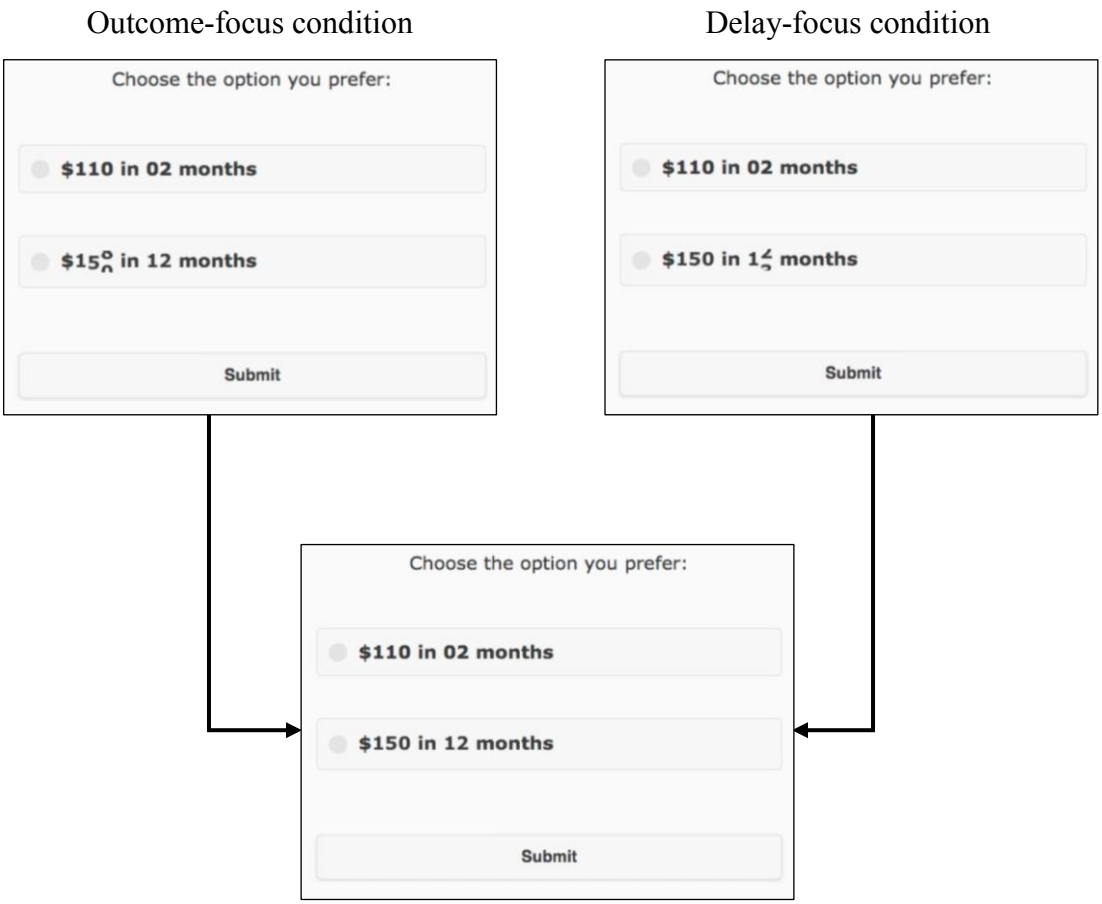

Figure 3.2. Example screenshots of the experimental procedure in Experiment 1. In the outcome-focus condition, each intertemporal choice (except the first item in each block) was preceded by a transition from a different LL outcome to the current one (1000 ms), while the environment was otherwise kept unchanged. In the delay-focus condition, each intertemporal choice (except the first item in a block) was preceded by a transition from another LL delay to the current one (1000 ms), while the environment was otherwise kept unchanged.

Procedures. Participants were randomly assigned to one of the two conditions. To further control the effect of the relative position of the two options, the vertical placement of the SS and the LL options were counterbalanced among participants. For half of the participants from each condition, SS was placed above LL; for the other half, SS was placed below LL. After the intertemporal choice task, participants completed the Cognitive Reflection Test (CRT; Frederick, 2005) and demographic questions including gender, age, ethnical identity, annual income, educational qualifications and employment status. Being loosely related to the core focus of the chapter, results related to CRT were reported in Appendix 3C.

Data analysis. Data analysis was carried out at the aggregate level. Binary data points from individual participants were aggregated into binomial data on each item in each condition. Mixed-effect models were applied to the aggregate data with a random intercept for each intertemporal choice item to capture the variation of the relative attractiveness of LL to SS across items.[20] Statistical model fitting consisted of two steps. First, the probability of choosing LL was predicted by mixed-effect models with a logit link function. Specifically, two models were used:

(1) $\hat{p}_i = \text{logit}^{-1}\left(\alpha_{j[i]} + \beta_1 x_i^{\text{Attention}}\right)$ and

(2) $\hat{p}_i = \text{logit}^{-1}\left(\alpha_{j[i]} + \beta_1 x_i^{\text{Attention}} + \beta_2 x_i^{\text{BackgroundContrast}} + \beta_3 x_i^{\text{Attention}} x_i^{\text{BackgroundContrast}}\right)$,

where $\alpha_{j[i]}$ (j = 1, …, 81) was the random intercept for each item. The random intercept was used to capture the relative attractiveness between SS and LL for each item. Model (1) had only attention, an effect-coded variable (-0.5 = delay-focus; 0.5 = outcome-focus), as the fixed-effect predictor while Model (2) involved both attention and background contrast. Second, the predicted probability of choosing LL, $\hat{p}_i$, was fitted to data with the binomial distribution $k_i \sim \text{Bin}(N_i, \hat{p}_i)$, where k_i was frequency of LL choices and N_i was the corresponding sample size. Results of the two models are shown in Table 3.1.

Bayesian modelling was used for parameter estimation and model evaluation (see Kruschke, 2010). To be conservative, the prior distribution used for all parameters was the standard normal distribution, rather than the uniform distribution. Estimation of parameters was summarized by the median (Md) and 95% high density interval (95% HDI) of posterior distributions. Model evaluation and comparisons were based on the Deviance Information Criterion (DIC; Gelman, Carlin, Stern, Dunson, Vehtari, & Rubin, 2014a; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002): The lower the DIC value, the better the model. The descriptive accuracy of some models was further checked with posterior predictive check (Gelman, Carlin, Stern, Dunson, Vehtari, & Rubin, 2014b).

Bayesian modelling and Deviance Information Criterion were approximated by Markov chain Monte Carlo (MCMC) in JAGS (Plummer, 2003), which was called

[20] We did not involve the formal intertemporal choice models (as in Chapter 2) to capture the variation across items because it is highly controversial which model is descriptively accurate (see also Cavagnaro et al., 2016a; Dai & Busemeyer, 2014; Read, 2001; Scholten et al., 2014).

by the *runjags* package (Denwood, in press) in R (R Core Team, 2016). Three independent Markov chains were run for each model. Each chain consisted of 10,000 burn-in samples, which served as the initial adaption period and were discarded from formal approximation, and 50,000 formal samples. Convergence among the three chains was checked with the Gelman-Rubin statistic (Gelman &Rubin, 1992). The three chains summed up to 150,000 samples from the posterior distribution.

3.4.2 Experiment 1: Results and discussion

Among all 191 participants, five did not choose the dominant option in at least seven out of the nine screening items (see the histogram in Figure 3.3 for a full distribution of the number of dominant choices among the nine screening items) and thus were excluded from the formal analysis, leaving us with data from 186 participants in the analysis. Note that the results were the same when all data were included in the following analysis.

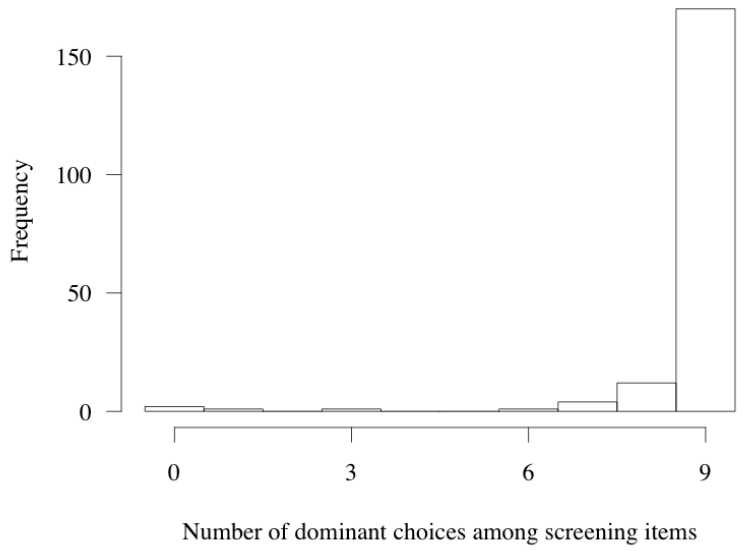

Figure 3.3. Histogram of the number of dominant choices of the nine screening items.

Attention effect. A preliminary check of the attention effect is by comparing the proportion of LL choices in the two experimental conditions. Participants in the outcome-focus condition made 40.1% LL choices while those in the delay-focus condition only chose 29.5% LL, showing a sizable attention effect on intertemporal choice. The Bayesian estimation of the attention effect in Model (1) confirmed that participants from the outcome-focus condition were much more likely to choose LL

than those from the delay-focus condition ($Md_{\beta 1}$ = 0.581, 95% $HDI_{\beta 1}$ = [0.506, 0.657]; see Table 3.1).

Background contrast effect. As explained earlier, this manipulation of attention would also change the background contrast at the item level as a by-product. Specifically, we distinguished between two types of background contrast: the global background contrast and the local background contrast. The *global background contrast* referred to all the tradeoffs *in the experiment* that have been exposed before the current tradeoff. Because of the randomization of both blocks and items within blocks, any two items would appear prior to or after each other with even odds in both conditions, so global contrast was counterbalanced between conditions at the aggregate level.

The *local background contrast* referred to all the tradeoff *in the same block* that have been exposed before the current tradeoff. Because the SS option was constant across items, we quantified the local background contrast with the relative rank of the LL option of all LL options in the same block, ranging from -1 (dominated by all other LL options), through 0, to 1 (dominating all other LL options).[21] Practically, it was calculated with the difference between the proportion of LL options that an LL option dominated and the proportion of LL options that dominated the LL proportion among all other items in the same block. For example, in the outcome-focus condition, $150 in 8 months dominated three LL options and were dominated by five LL options in the same block. Its local background contrast is quantified as 3/8 – 5/8 = -0.25. In the delay-focus condition, $150 in 8 months dominated six LL options and were dominated by two LL options in the same block. Its local background contrast is quantified as 6/8 – 2/8 = 0.5.

Model (2) involved the quantified local background contrast and thus statistically controlled the background contrast effect. The result from this model was consistent with Model (1) in terms the attention effect on intertemporal choice ($Md_{\beta 1}$ = 0.563, 95% $HDI_{\beta 1}$ = [0.483, 0.641]; see Table 3.1). In addition, it revealed a strong background contrast effect in the predicted direction: The more dominated LL options (or the less dominating LL options) there were in the same block, the more likely the

[21] Theoretically, only the items that appeared before the current one is the "background". Because the order of items in a block is randomized, each item has even odd to appear before or after another item. So when data are aggregated across participants, the overall pattern should be a good approximation of the aggregate "background" for the current item.

current LL is to be chosen ($Md_{\beta 2}$ = 0.484, 95% $HDI_{\beta 2}$ = [0.404, 0.564]). Local contrast did not influence the attention effect ($Md_{\beta 2}$ = 0.044, 95% $HDI_{\beta 2}$ = [-0.146, 0.235]).

Table 3.1 Model evaluation and parameter estimation (Experiment 1)

Fixed-effect	Parameter	Estimation	
		(1)	(2)
Attention	β_1	**0.581 [0.506, 0.657]**	**0.563 [0.483, 0.641]**
Background Contrast	β_2		**0.484 [0.404, 0.564]**
Attention × Background Contrast	β_3		0.044 [-0.146, 0.235]
DIC		1036.535	930.558

Note. Attention is an effect-coded variable: -0.5 = delay-focus; 0.5 = outcome-focus. *Background contrast* ranges from -1 to 1, by an increment of 0.25. The estimates outside of the brackets are the median (Md) and the estimates in the brackets are the 95% High Density Intervals (95% HDIs) of the 150,000 samples from the posterior distribution. The 95% HDIs of the cells in boldface do not cross 0.

Involving the local background contrast in the model increased the performance of the model. According to DIC, Model (2), which involved background contrast as a fixed-effect predictor along with attention, performed better than Model (1), which involved attention as the only fixed-effect predictor (DIC_{M1} = 1036.535; DIC_{M2} = 930.558). To further check the goodness of fit of Model (2) to data, we ran a posterior predictive check by comparing the predicted probability of choosing LL based on posterior distributions of parameters with the original proportion of LL choice (Figure 3.4). We found that Model (2) described the data in Experiment 1 very well. Ninety-eight percent (159 out of 162) of the original LL proportions from data are accommodated within the predicted 95% HDIs based on the posterior distribution of parameters.

To summarise, Experiment 1 showed a strong attention effect on intertemporal choice, with an attention-driven shift of preference of 10.6% LL choices (40.1% vs. 29.5%). The robustness of this effect was confirmed by two mixed-effect Bayesian models (Models 1-2 in Table 3.1). The finding was consistent with both attribute-wise and component-wise attention effects on intertemporal choice, which predicted that

attending to the LL outcome, relative to attending to the LL delay, reduced impatience, (see Figure 3.1b-c).

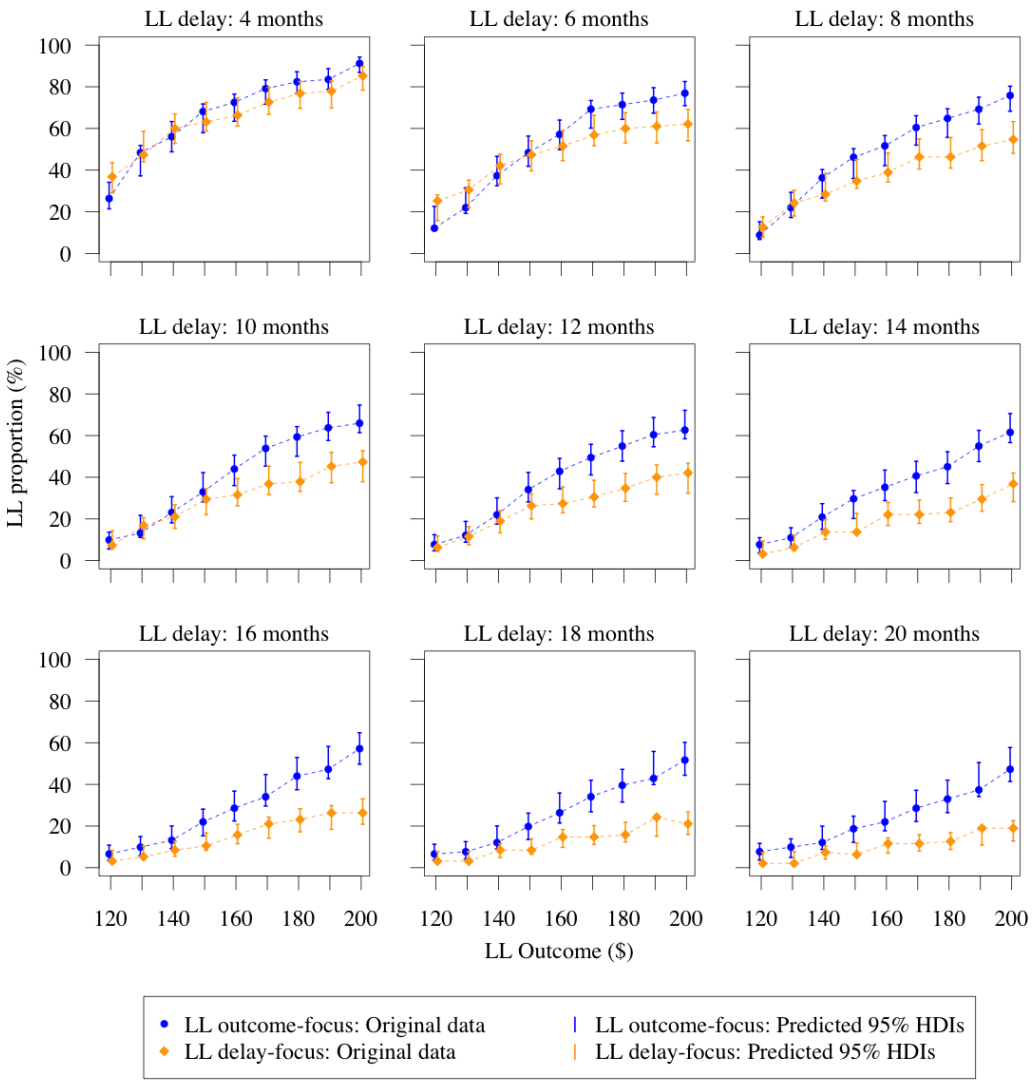

Figure 3.4. Posterior predictive check of Model (2) for Experiment 1. The dots represent original proportions of LL choices and the vertical lines represent the 95% high density intervals (HDIs) of the predicted probability of choosing LL based on the 150,000 samples drawn from posterior distributions.

3.5 Experiment 2: Varying SS Components

Experiment 2 disentangled the attribute-wise and the component-wise attention effects. It was similar to Experiment 1 except that attention was manipulated on the smaller-sooner (SS) option. Thus, in the *outcome-focus* condition, attention was directed to the SS outcome; In the *delay-focus* condition, attention was directed to the

SS delay. The attribute-wise attention effect predicted that focusing attention on the SS outcome (outcome-focus condition), invoking the comparison of options along the outcome attribute, made people less impatient than attending to the SS delay (delay-focus condition), invoking the comparison along the delay attribute (Figure 3.1b). By contrast, the component-wise effect predicted that attending to the SS outcome made SS more attractive and attending to the SS delay made SS less attractive. So, the component-wise effect predicted that attending to the SS outcome would make people more impatient than attending to the SS delay (Figure 3.1c).

3.5.1 Experiment 2: Methods

Participants. Participants were recruited using the same procedure as in Experiment 1. There were 196 USA residents recruited via Prolific Academic (76 females and 120 males). They were on average 31.0 years old. 62.8% of them had completed a university degree or higher. After completing the experiment, each of them received a flat payment of $1.

Intertemporal-choice task. Each participant made 90 choices, of which 81 were target items and 9 were screening items with a dominant option. LL was constant across all items, which was receiving $190 in 20 months while SS varied in both the magnitude of the outcome and the length of the delay. The SS outcomes of the target items could take nine values from $100 to $180 by an increment of $10. The SS delays of the target items could take nine different values from 2 months to 18 months by an increment of 2 months.

Design. There were two conditions, manipulated between participants. In the *outcome-focus* condition, the 81 target items were divided into nine blocks so that only one of the nine SS delays, but all SS outcomes, appeared within a block. In each block, an additional screening item was added, of which the SS outcome was $200 and the SS delay was the same as the delay in the corresponding block (e.g., $200 in 6 months), making LL dominated by SS. In the *delay-focus* condition, items were also divided into nine blocks. In each block, all SS delays, but only one SS outcome, appeared. The SS delay of the screening item was 20 months and the SS outcome was the same as the SS outcome of target items in the corresponding block, making SS dominated by LL. In both conditions, the substitution of the varying value across items in a block

was accompanied with a one-second inter-item transition animation, as in Experiment 1.[22]

Procedures. Participants were randomly assigned to one of the two conditions. Within each condition, the order of blocks was randomized and the order of items in each block was randomized across participants. The relative position of SS and LL were counterbalanced across participants. For half of the participants from each condition, SS was always above LL; for the other half, SS was always below LL. After the intertemporal choice task, participants completed the Cognitive Reflection Test (CRT) and demographic questions as in Experiment 1 (see Appendix 3A-B for details; see Appendix 3C for results relating to CRT).

Data analysis. As in Experiment 1, parameter estimation was carried out with mixed-effect Bayesian modelling. Different fixed-effect terms were used for models as shown in Table 3.2 and random intercepts were applied to each item to capture the relative attractiveness of LL to SS across items. The Bayesian analysis was approximated with the same MCMC simulation procedure as in Experiment 1. Model evaluation and comparison were based on DIC and posterior predictive check.

3.5.2 Experiment 2: Results and discussion

Data from three participants were excluded from the formal analysis because they did not choose the dominant option in at least seven out of the nine screening items, leaving data from 193 participants in the formal analysis.

Attention effect. Participants in the outcome-focus condition chose 41.2% LL and those in the delay-focus condition chose 43.0% LL, showing a slight attention effect on intertemporal choice. The direction of the effect was consistent with the component-wise effect, but was inconsistent with the attribute-wise effect. This effect was confirmed by the mixed-effect Bayesian model with manipulated attention as the fixed-effect predictor (Model (1) in Table 3.2). Participants in the outcome-focus condition were less likely to choose LL than those from the delay-focus condition ($Md_{\beta 1}$ = -0.107, 95% $HDI_{\beta 1}$ = [-0.184, -0.033]).

[22] The link to Experiment 2 is: http://gilum.lnx.warwick.ac.uk/wbs/experiment2/

Table 3.2 Model evaluation and parameter estimation (Experiment 2).

Fixed-effect	Parameter	Estimation	
		(1)	(2)
Attention	β_1	**-0.107 [-0.184, -0.033]**	**-0.152 [-0.230, -0.076]**
Background Contrast	β_2		**-0.439 [-0.517, -0.359]**
Attention × Background Contrast	β_3		**-0.272 [-0.474, -0.067]**
DIC		1020.769	930.206

Note. Attention is an effect-coded variable: -0.5 = delay-focus; 0.5 = outcome-focus. *Background contrast* ranges from -1 to 1, by an increment of 0.25. Model evaluation and parameter estimation (Experiment 2): The estimates outside the brackets are the median (Md) and the estimates in the brackets are the 95% High Density Intervals (95% HDIs) of the 150,000 samples from the posterior distribution. The 95% HDIs of the cells in boldface do not cross 0.

Background contrast effect. The local background contrast was quantified as in Experiment 1, ranging from -1 (the current SS option was dominated by all others in the same block) to 1 (the current SS option dominated all others in the same block). Model (2) in Table 3.2 controlled the background contrast effect. In Model (2), the direction of the attention effect was consistent with that in Model (1) ($Md_{\beta 1}$ = -0.152, 95% $HDI_{\beta 1}$ = [-0.230, -0.076]). The local background contrast also influenced intertemporal choice in the predicted direction ($Md_{\beta 2}$ = -0.439, 95% $HDI_{\beta 2}$ = [-0.517, -0.359]). There was an interaction between the attention effect and the background contrast effect on intertemporal choice ($Md_{\beta 3}$ = -0.272, 95% $HDI_{\beta 3}$ = [-0.474, -0.0067]), which means the attention effect was moderated by the local background contrast. This moderation effect, however, did not attenuate the attention effect as shown in the main effect of attention in Model (2).

Involving background contrast improved the performance of the model. As shown in Table 3.2, the DIC value of Model (2) was smaller than the DIC value of Model (1) (DIC_{M1} = 1020.769; DIC_{M2} = 930.206). The goodness of fit of Model (2) to the data in Experiment 2 was further confirmed by a posterior predictive check. Comparisons between the original data and the posterior predictions suggested that

98.8% (160 out of 162) original LL proportions points were within the corresponding 95% HDIs of the predicted probability of choosing LL based on the 150,000 samples from the posterior distribution of parameters (see Figure 3.5).

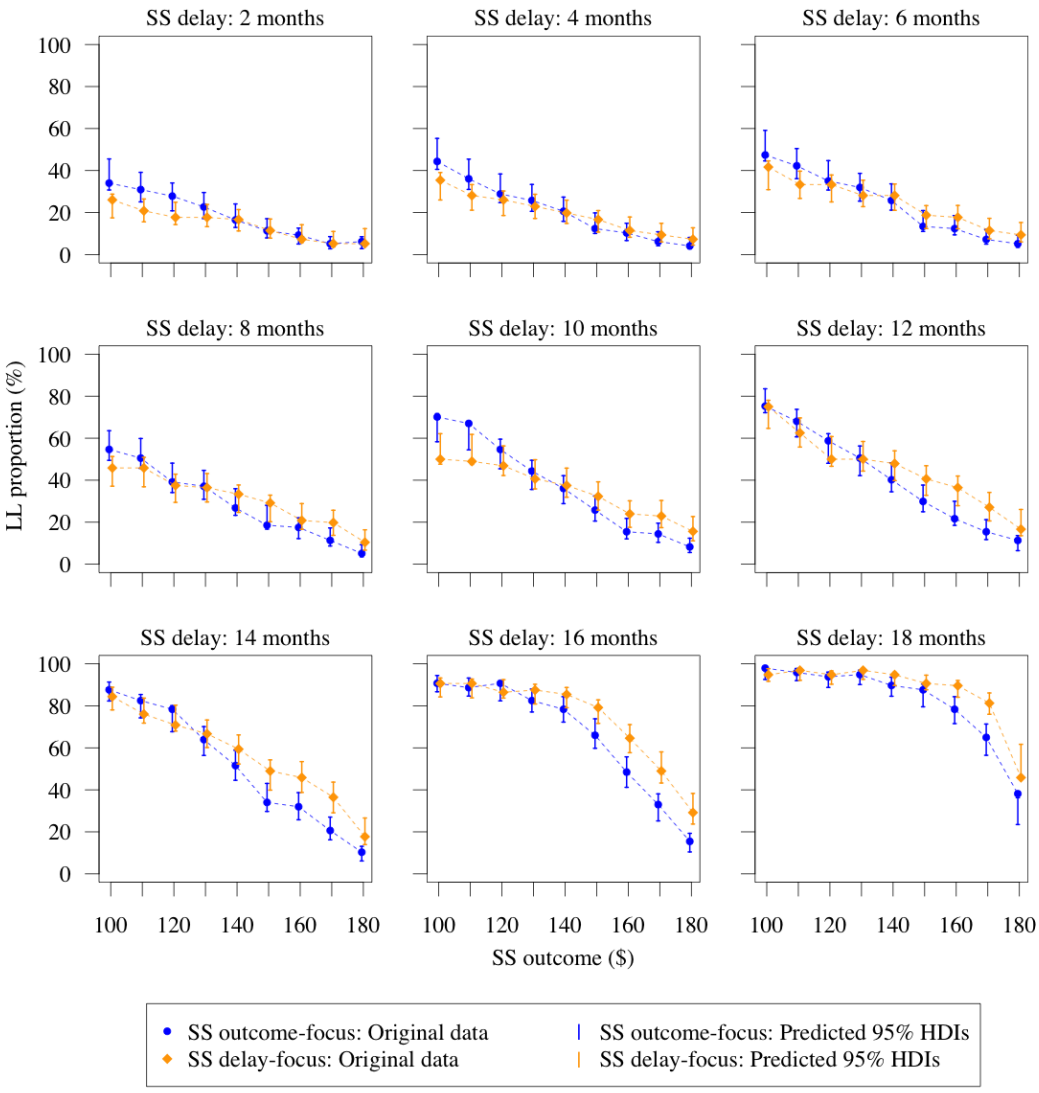

Figure 3.5. Posterior predictive check of Model (2) for Experiment 2. The dots represent original proportions of LL choices and the vertical lines represent the 95% high density intervals (HDIs) of the predicted probability of choosing LL based on the 150,000 samples drawn from posterior distributions.

Comparisons between Experiments 1 and 2. Although the direction of the attention effect was inconsistent with prediction of the attribute-wise effect, it did not rule out the possibility that the attribute-wise effect co-exist with the component-wise effect. Indeed, it was very likely that they co-existed because the attention-driven

preferential shift in in Experiment 2, where the attribute-wise and the component-wise effects offset each other, was much smaller than that in Experiment 1, where the two ways of the attention effects compensated each other. This contrast had two manifestations. First, between-condition difference in the average proportion of LL choice of all items was 10.6% in Experiment 1, which was much larger than 1.8% as in Experiment 2 (Figure 3.6a). Using a frequentist independent-sample t-test, the difference reached significance at $\alpha = .001$ ($t_{155} = 6.48$, $p < .001$). This difference could not be attributed to a floor or ceiling effect in Experiment 2 because the average proportion of LL choice in Experiment 2 is even closer to 50-50 than that in Experiment 1. Thus, there should be no floor or ceiling in Experiment 2 that curbed the effect size. Second, the Bayesian estimation of the attention effect (β_1), which could be exponentiated to Odd Ratio as the effect size, from model fitting was much larger in Experiment 1 than in Experiment 2 (Figure 3.6b). Regardless of which model was used for the estimation, the estimated effect size was constantly larger in Experiment 1 than in Experiment 2.

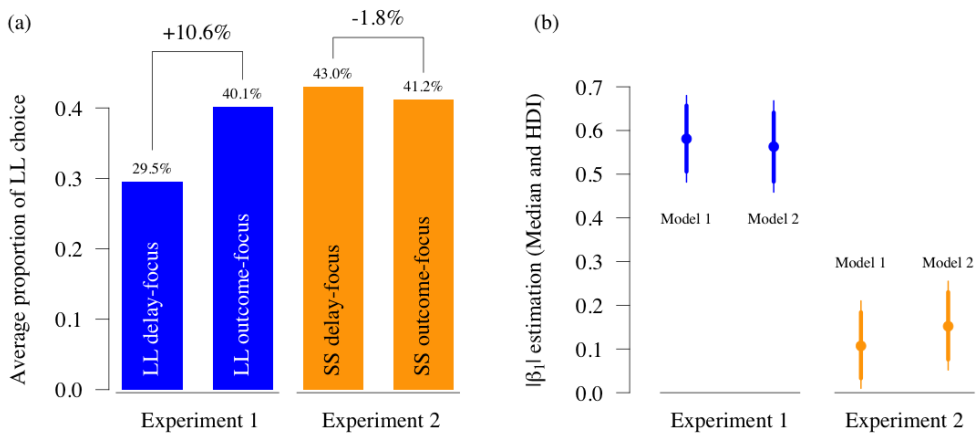

Figure 3.6. Comparisons of the two experiments. (a) The between-condition difference of the overall proportion of LL choice is larger in Experiment 1 (blue bars), where attribute-wise and component-wise effects compensate each other, than in Experiment 2 (orange bars), where they offset each other. (b) Bayesian estimation of the size of the attention effect (β_1) is larger in Experiment 1 (blue lines) than in Experiment 2 (orange lines). The points represent median values of posterior distributions and the lines 95% and 99% HDIs respectively. Note that the direction of the effect in Experiment 2 was mirrored for the ease of comparing the effect sizes.

To sum up, when the attribute-wise and component-wise effects were disentangled, the direction of the net attention effect found in Experiment 2 was not consistent with the attribute-wise effect, neither the option-wise effect. Thus, it was direct evidence of the component-wise attention effect on intertemporal choice. Moreover, the findings did not rule out the possibility that the attribute-wise attention effect co-existed with the component-wise attention effect. Comparisons of the effect sizes from Experiment 1 and Experiment 2 suggested that the co-existence of the attribute-wise and the component-wise attention effects could explain the gap in the effect sizes from the two experiments better than the absence of the attribute-wise attention effect in the two experiments.

3.6 Discussion

Beyond the option-wise attention effect, this study provided evidence for two ways of the attention effects (i.e., attribute-wise and component-wise attention effects) on intertemporal choice, using a new method to manipulate attention. In Experiment 1, manipulated attention produced a sizable attention effect on intertemporal choice when the attribute-wise and component-wise effect compensated each other. Experiment 2 further provided direct evidence for the component-wise attention effect, when the two effects were disentangled. Finally, the large gap between the attention effect sizes in the two experiments suggested a probable co-existence of the attribute-wise and the component-wise attention effects.

This study also generalise some previous findings. Most previous studies have found evidence for the option-wise attention effect on value-based decisions, but these effects were based on strong assumptions of the decision process (e.g., Fiedler & Glőckner, 2012; Franco-Watkins et al., 2016; Krajbich & Rangel, 2011; Krajbich et al., 2010; Stewart et al., 2016). When studying the option-wise attention effect only, researchers were also, at least implicitly, assuming the alternative-based evaluation rule for decision making. While many studies suggest attribute-based evaluation rule for decision making (e.g., Arieli et al., 2011; Dai & Busemeyer, 2014; Scholten & Read, 2010; Tversky, 1972; as well as Chapter 2 in this thesis), the investigation of the attribute-wise attention effect is rare (see Fisher & Rangel, 2014, Experiment 2; Hare et al., 2011). In face of both lines of evidence, the present study did not make a particular assumption of the decision process and the way of the attention effect and found evidence for the co-existence of multiple ways of the attention effect. Moreover,

with the (indirect) evidence for the attribute-wise attention effect, the results also suggest the existence of attribute-based evaluation rule for intertemporal choice.

This study has several limitations. First, although we are pretty confident that this attention manipulation was valid, we lacked the process data, such as eye-tracking, to prove that our participants were paying more attention to the component of interest, which we intended to direct their attention to, than others. Second, the results did not provide direct evidence for the attribute-wise attention effect. However, because of the similarity between the two experiments, we were able to make an inter-experiment comparison and thus inferred that the attribute-wise attention effect co-existed with the component-wise attention effect Third, in this study, we did not study the most studied option-wise attention effect.

To conclude, this study showed that attention, which was manipulated by simply varying different components across items, alters intertemporal choice, thus driving intertemporal preference. The results suggested that, beyond the option-wise effect, attention could operate on an attribute and a component individually, therefore influencing choice behaviour differently. In addition, the experiments also found robust background contrast effects, suggesting violations of sequential independence in intertemporal choice. Collectively, these results suggested that the elicited intertemporal preference was highly uncertain, contingent on attention allocation, as well as background contrast. While the elicitation of time preference inevitably involves varying one or another component across items, a better understanding of how attention drives intertemporal choice is especially valuable to de-biasing the attention effect that is very likely to emerge in the elicitation procedure.

CHAPTER 4 DETECTED PATTERNS OF IMPATIENCE

METHODOLOGICAL CONCERNS

Standard models of intertemporal discounting assume stationarity or *constant impatience* (CI) over time (Fisher, 1930; Fishburn & Rubinstein, 1982; Koopmans, 1960; Samuelson, 1937). Despite the normative appeal and mathematical tractability of constant impatience, studies have documented an array of phenomena that are anomalous to constant discounting (see Frederick, Loewenstein, & O'Donoghue, 2002; Read et al., in press for further reviews). One of the most prominent phenomena is the *delay effect*, which means that people discount at a higher rate for short delays than for long ones. For example, someone indifferent between $1,000 in one year and $800 today (implying a discount rate of 20% per annum) would probably prefer $1250 in two years to $800 today (implying a discount rate lower than 20% per annum).

The delay effect is one of the most robust phenomena anomalous to constant discounting and has been corroborated in a host of studies (e.g., Ainslie, 1975; Benzion et al., 1989; Bleichrodt & Johannesson, 2001; Cairns & van der Pol, 2000; Green et al., 1994b; Kirby & Marakovic, 1995; Pender, 1996; Myerson & Green, 1995; Read & Read, 2004; Richards, Zhang, Mitchell, & de Wit, 1999; Rodriguez & Logue, 1988; Thaler, 1981). Based primarily on the delay effect, many researchers proposed that the discount rate declined over time, which was called *decreasing impatience* (DI). In the above example, with the inferred discount rate of 20% per annum over the first year from the indifference between $1,000 in one year and $800 today, the preference of $1250 in two years to $800 today suggests that the discount rate over the second year of the longer delay is lower than 20% per annum.

However, the delay effect confounds decreasing impatience with subadditive discounting (Read, 2001). Subadditive discounting means that people are more impatient when a delay is divided into sub-intervals than when the delay is undivided. This effect has been replicated in many studies (e.g., Kinari, Ohtake, & Tsutsui, 2009; McAlvanah, 2010; Read, 2001; Read & Roelofsma, 2003; Scholten & Read, 2006; 2010). Reconsider the example of the delay effect above. The preference of $1250 in two years to $800 today can be attributed to either a decrease of discount rate over the

second year (i.e., decreasing impatience) or a decrease of discount rate over the undivided delay of two years than over the two consecutive one-year sub-intervals (subadditive discounting). Thus, the delay effect is not convincing evidence for decreasing impatience.

The pattern of impatience (sometimes described as constant vs. hyperbolic vs. anti-hyperbolic discounting) has thereafter been debated for decades in the study of intertemporal choice. Many studies attempted to examine the pattern of impatience while subadditive discounting was controlled. Unsurprisingly, testing at the individual level often showed heterogeneous patterns, with some decreasingly impatient and other increasingly impatient (e.g., Attema, Bleichrodt, Rohde, & Wakker; 2010; Augenblick, Niederle, &Sprenger, 2015; Bleichrodt, Gao, &Rohde, 2016; Halevy, 2015; Olea & Strzalecki, 2014). Still, testing at the aggregate level also resulted in a mixture of constant impatience (CI; e.g., Halevy, 2015; Read, 2001), decreasing impatience (DI; e.g., Bleichrodt et al., 2016; Chark, Chew, & Zhong, 2015; Green, Myerson, Macaux, 2005; Keren & Roelosfma, 1995; Kinari, Ohtake, & Tsutsui, 2009; Read & Read, 2004; Read & Roelofsma, 2003; Scholten & Read, 2006; 2010; Sopher & Sheth, 2006; Weber & Chapman, 2005) and increasing impatience (II; e.g., Attema et al., 2010; Read, Olivola, & Hardisty, 2016; Sayman & Öncüler, 2009; Scholten & Read, 2006; 2010). Increasing impatience is the reversal of decreasing impatience, which means that people discount more for remote future than for close future.

4.1 A Design Bias

Among all, a vast majority of the studies used pairs of intertemporal choice items between smaller sooner (SS) and larger-later (LL) options to detect the pattern of impatience at the aggregate level (e.g., Chark, Chew, & Zhong, 2015; Holcomb & Nelson, 1992; Keren & Roelosfma, 1995; Read, Olivola, & Hardisty, 2016; Sayman & Öncüler, 2009; Scholten & Read, 2006; 2010; Sopher & Sheth, 2006; Weber & Chapman, 2005). For example, in the pair of choices below,

Choice 1a: $100 today (A) or $110 in three months (B)

Choice 1b: $100 in three months (C) or $110 in six months (D),

if someone chooses A and D, she is showing decreasing impatience. By contrast, if someone chooses B and C, she is showing increasing impatience. However, the patterns of those choosing A and C [or B and D] cannot be identified. Thus, studies employing such paradigms made an implicit assumption that the unobserved patterns

were homogeneous to the observed and thus generalized from the observed patterns to the unobserved.

However, to our best knowledge, the assumption that the unobserved patterns of impatience are homogeneous to the observed patterns has not been empirically tested. If this assumption does not hold, we argued that the detected aggregate pattern of impatience could be biased by the design of pairs. To illustrate, in the pair of Choice1a and 1b, because the LL options are relatively less attractive than the SS options, impatient people are likely to choose A and C regardless of their patterns of impatience and thus their patterns are unobservable. Thus, this pair is likely to identify only the patterns of impatience from patient individuals. By contrast, in the pair of Choice 2a and 2b below:

Choice 2a: $100 today (E) or $300 in three months (F)

Choice 2b: $100 in three months (G) or $300 in six months (H),

because the LL options are relatively more attractive than the SS options, patient individuals are very likely to choose F and H regardless of their patterns of impatience. Correspondingly, the pair of Choice 2a and 2b is likely to identify only the patterns of impatience from impatient individuals.

As illustrated above, a testable hypothesis from the homogeneity assumption is the independence between individuals' pattern of impatience and their degree of impatience. Suppose that impatient individuals are more likely to exhibit decreasing impatience (DI) than patient ones and patient individuals are more likely to exhibit increasing impatience (II) than impatient ones. Then the pair of Choice 1a and 1b is likely to show increasing impatience, but the pair of Choice 2a and 2b is likely to show decreasing impatience, at the aggregate level.

4.2 An Order Effect

Studies that investigate discounting using either the switch paradigm or other paradigms ask two matched sets of intertemporal choice questions: we will call these a *base* set and a *delayed* set. As in the examples above, a common front-end delay (FED) is added to the delays of the base set to form the delayed set. In many studies both sets are presented in random order (e.g., Bleichrodt et al., 2016; Green et al., 2005; Holcomb & Nelson, 1992; Kinari, Ohtake, & Tsutsui, 2009; Read, 2001; Read & Read, 2004; Read & Roelofsma, 2003; Read et al., 2016; Scholten & Read, 2006; Sopher & Sheth, 2006), in others the two sets are presented separately and sequentially in order. For example, Sayman and Öncüler (2009, Study 2a) presented the base set

prior to the delayed set to half of their participants and presented the delayed set prior to the base set to the other half, but no comparison was made in the paper. Chark et al. (2015) used one base set and two delayed sets and the base set was presented at the top of a list of choices, followed by the two delayed sets. Some studies assigned the different sets to different groups of participants (Keren & Roelofsma, 1995; Sayman & Öncüler, 2009, Study 2b; Weber & Chapman, 2005).

Does the presentation order influence the identified discount pattern? In behavioural decision research, order effects have been observed in financial decision making (Mellers, Schwartz, Ho, & Ritov, 1997; Thaler, Tversky, Kahneman, & Schwartz, 1997), social choice (Knez & Camerer, 2000; Vlaev & Chater, 2007) and intertemporal choice (Dai, Grace, & Kemp, 2009). A well-documented order effect is the background contrast effect, which suggests that exposure to prior tradeoff will influence their choice in the current tradeoff (Priester et al., 2004; Simonson & Tversky, 1992; Stewart et al., 2003; 2015; Ungemach et al., 2011; Vlaev et al., 2007; 2009; Walasek & Stewart, 2015). However, to our best knowledge, no study has directly investigated the order effect on the identified discount pattern. Given that different studies presented intertemporal choice questions with different orders, if there were an order effect on the identified discount pattern, it could contribute to conflicting evidence on hyperbolic discounting.

4.3 The Present Study

This study focused on the methodological issues on the pattern of impatience and has three goals. The first was to study the relationship between individuals' pattern of impatience and their degree of impatience, therefore examining a potential design bias when pairs of items were used to detect the aggregate pattern of impatience. The second goal was to test an order effect on the detected pattern of impatience. The two contributions not only filled the gaps in the literature, but could also contribute to the understanding of the heterogeneous and conflicting findings regarding the pattern of impatience in the literature. Third, it could also detect the pattern of impatience in various conditions.

Two experiments were conducted for this study. Both experiments adapted Kirby, Petry and Bickel's (1999) Monetary Choice Questionnaire to measure the degree of impatience and to detect the pattern of impatience. The original Monetary Choice Questionnaire had 27 intertemporal choice items, of which each was a binary choice between an immediately available smaller-sooner (SS) option and a delayed

larger-later (LL) option (see Table 4.1). To detect the pattern of impatience, another 27 items were created by adding a front-end delay of 100 days to each of the 27 original intertemporal choice items. A comparison between the new items and the original items enabled us to identify the pattern and, to some extent, to quantify the degree of decreasing impatience.

The advantages to use the adapted Monetary Choice Questionnaire were twofold. First, this questionnaire was easy to use and time-saving. With only 27 original items and 27 adapted items, participants mostly completed the task within ten minutes. Second, the 27 pairs of choices offered extensive variation in terms of the relative attractiveness between SS and LL, which enables us to identify the aggregate pattern of impatience with different subsets of participants along the spectrum of the relative attractiveness between SS and LL (see Figure 4.2 showing that the average SS proportion per pair ranges from one end [0%] to the other [100%]).

Table 4.1 Intertemporal choice items and implied hyperbolic discount rates in both the no-FED (front-end delay = 0) condition and the FED (front-end delay = 100) condition.

Pair ID	Smaller amount (x_L)	Larger amount (x_S)	Interval (t)	Magnitude level	Implied k (no-FED)	Implied k (FED)
1	11	30	7	Small	0.24675	**-0.01042**
2	15	35	13	Small	0.10256	**-0.01108**
3	14	25	19	Small	0.04135	**-0.01319**
4	24	35	29	Small	0.01580	**-0.02723**
5	19	25	53	Small	0.00596	0.01474
6	25	30	80	Small	0.00250	0.00333
7	22	25	136	Small	0.00100	0.00111
8	28	30	179	Small	0.00040	0.00042
9	34	35	186	Small	0.00016	0.00016
10	20	55	7	Medium	0.25000	**-0.01042**
11	25	60	14	Medium	0.10000	**-0.01111**
12	27	50	21	Medium	0.04056	**-0.01327**
13	34	50	30	Medium	0.01569	**-0.02759**
14	40	55	62	Medium	0.00605	0.01531
15	49	60	89	Medium	0.00252	0.00337
16	54	60	111	Medium	0.00100	0.00111
17	47	50	160	Medium	0.00040	0.00042
18	54	55	117	Medium	0.00016	0.00016
19	31	85	7	Large	0.24885	**-0.01042**
20	33	80	14	Large	0.10173	**-0.01109**
21	41	75	20	Large	0.04146	**-0.01318**
22	54	80	30	Large	0.01605	**-0.02653**
23	55	75	61	Large	0.00596	0.01476
24	69	85	91	Large	0.00255	0.00342
25	67	75	119	Large	0.00100	0.00112
26	80	85	157	Large	0.00040	0.00041
27	78	80	162	Large	0.00016	0.00016

4.4 Experiment 1: Design Bias

In Experiment 1, we primarily tested the correlation between participants' degree of impatience and their pattern of impatience. The degree and pattern of impatience were simultaneously measured with the adapted Monetary Choice Questionnaire as described above. In addition, the Cognitive Reflection Test (CRT; Frederick, 2005) and the Consideration for Future Consequences scale (CFC; Strathman, Gleicher, Boninger, & Edwards, 1994) were also administrated (See Appendices 3A and 4A for the details of the questionnaires). By testing the relationship between the degree and the pattern of impatience, Experiment 1 could examine a design bias when pairs of intertemporal choice items were used to detect the aggregate pattern of impatience.

4.4.1 Experiment 1: Methods

Participants. Participants for Experiment 1 were 104 adults (65 females, 29 males and others unidentified) from the UK, recruited online via prolific academic (https://prolific.ac/). They were on average 37.87 (SD = 11.86) years old.

Design. Experiment 1 used a within-subject design. Two conditions were involved: The no-FED condition and the FED condition. Using a within-participant design, Experiment 1 could identify whether an individual exhibited decreasing, increasing or constant impatience in intertemporal choice at both individual and aggregate levels.

Tasks and procedures. Participants were asked to complete three tasks: the intertemporal choice task, the Cognitive Reflection Test (CRT) and the Consideration for Future Consequences (CFC) scale. As mentioned earlier, the intertemporal choice task was adapted from the Monetary Choice Questionnaire. The no-FED condition used the original 27 items from the Monetary Choice Questionnaire (e.g., choosing between receiving £11 today and receiving £30 in seven days). In the FED condition, a front-end delay of 100 days was added to each of the original items from the Monetary Choice Questionnaire (e.g., choosing between receiving £11 in 100 days and receiving £30 in 107 days). In addition, three screening items, of which each had a dominant option, were added to each condition to check if the participants were just choosing randomly (six screening items in total; see Appendix 4B for details of the screening items). So, there were in total 60 intertemporal choice items. Items were presented individually to participants in a randomized order.

After the intertemporal choice task, participants were given the Cognitive Reflection Test (CRT) and the Consideration for Future Consequences (CFC) scale. CRT had three questions of which each was a four-option multiple choice. Previous studies showed that high cognitive reflection, measured with the CRT, reduced impatience (e.g., Frederick, 2005; see also Appendix 3C for Chapter 3). The CFC scale was a self-reported scale on to what degree people cared about consequences happening now or in the future. It had fourteen items, each of which was rated on a five-point Likert scale. The scale was slightly revised from the original version to increase its readability. Demographic information was recorded before participants completed the study (see Appendix 4C for details).

4.4.2 Experiment 1: Results and discussion

Data were firstly filtered by the screening items. Among all, only three participants selected the dominated option in only one of the six screening items. Thus, no data was eliminated from the formal analysis.

Model-free measure of impatience. Individual participants' degrees of impatience were quantified with the proportion of SS choices of all the 54 target items as in Myerson, Baumann and Green (2014). The larger the SS proportion, the more impatient the participant was. Quantifying the degree of impatience this way had three advantages. First, it was model-free and sidestepped the controversy on an appropriate model for intertemporal choice (see Doyle, 2012 for a collection of them). Second, this measure was straightforward and easily tractable in the FED condition, compared to the widely-used one-parameter hyperbolic discount function (Herrnstein, 1981; Mazur, 1987): $d(t)=\frac{1}{1+kt}$, where k is the hyperbolic discount rate. The hyperbolic discount function was the default model used to quantify the degree of impatience in the Monetary Choice Questionnaire (Kirby et al., 1999). However, it has a theoretical limitation in interpreting choices between two delayed outcomes because it produces nonsensical (negative) implied discount rates for 12 out of the 27 items in the FED condition (see the numbers in boldface in Table 4.1).

Third, the proportion measure of the degree of impatience could capture the information from the inferred hyperbolic discount rate, k, for the no-FED condition, in both a theoretical and an empirical sense. The Monetary Choice Questionnaire was designed to measure time preference by a two-way titration procedure, which was perfectly shown by the Spearman rank correlation coefficients between the interval

between SS and LL and the absolute/proportional amount differences (as shown in Table 4.2). With such an item structure, any intertemporal choice model assuming positive time preference should predict similar order of switch points. For example, any model, given any set of parameters, that predicts LL is preferred to SS in item #3, will also predict that LL is preferred to SS in items #4-9. This lay the theoretical basis to use the proportion of SS choices in the (adapted) Monetary Choice Questionnaire as the measure of impatience. Empirically, Myerson et al. (2014) found that the correlation between the proportion of SS choices and the logarithm of *k*, the estimated hyperbolic discount rate, in the Monetary Choice Questionnaire was extremely strong (Pearson correlation coefficient r = .97). Similarly, the data in Experiment 1 also revealed an extremely strong correlation between the logarithm of the estimated *k* and the proportion measure of impatience in the no-FED condition (Pearson correlation coefficient r = .94).

Table 4.2 Spearman correlation coefficients between the interval (*t*) and the absolute/proportional amount differences in the original Monetary Choice Questionnaire.

	Correlation coefficient between the interval (*t*) and	
	Absolute amount difference	Proportional amount difference
Small	-0.98	-1.00
Medium	-0.98	-0.98
Large	-1.00	-1.00
All	-0.87	-0.96

Note. The absolute amount difference is x_L - x_S. The proportional amount difference is $(x_L - x_S)/x_L$.

Decreasing impatience. The pattern of impatience could be qualitatively identified as decreasing impatience (DI), constant impatience (CI) and increasing impatience (II), and could be quantified along a continuum from extremely decreasing impatience, through constant impatience, to extremely increasing impatience. For brevity this quantification of the pattern will henceforth be called the *degree of decreasing impatience*. Like the measure of impatience, the degree of decreasing

impatience was quantified with a model-free measure: *d*, the number of SS choices in the FED condition minus the number of SS choices in the no-FED condition. Thus, a negative *d* meant decreasing impatience and a positive *d* meant increasing impatience. The smaller the *d* (< 0) value, the higher the degree of decreasing impatience. The larger the *d* (> 0) value, the higher the degree of increasing impatience.

As shown in Figure 4.1a, a majority of participants exhibited decreasing impatience (59.6%) while only a small proportion exhibited increasing impatience (24.0%) or constant impatience (16.3%). Thus, at the individual level, Experiment 1 revealed an overall pattern of decreasing impatience, $t(103) = 5.07$, $p < .001$.

Table 4.3 Data from an example pair of items showing decreasing impatience.

		FED condition	
		SS: £27 in 100 days	LL: £50 in 121 days
No-FED condition	SS: £27 today	8	14
	LL: £50 in 21 days	1	81

Note. In this pair of choices, only one participant shows increasing impatience by switching from LL in the no-FED condition to SS in the FED condition, while 14 exhibit decreasing impatience by switching from SS in the no-FED condition to LL in the FED condition. A McNemar's chi-square test on the off-diagonal suggests that this pattern is statistically significant, $\chi^2 = 11.27$, $p < .001$.

Another view at the aggregate level also revealed a general pattern of decreasing impatience. The 27 pairs of items across conditions allowed us to identify the pattern of impatience at the aggregate level. McNemar's chi-square tests showed that seven out of the 27 pairs exhibited decreasing impatience (*p*s < .05), while none of them demonstrated increasing impatience (see the left column in Figure 4.2). Table 4.3 gives an example of these pairs. In this example, fourteen participants switched from a choice of SS in the no-FED condition to a choice of LL in the FED condition, while only one made the reverse pattern of choices. McNemar's chi-square test on the off-diagonal confirmed the pattern of decreasing impatience shown in this item ($\chi^2 = 11.27$, $p < .001$). However, despite statistical significance, this specific pair only identified the switching patterns from only 15 out of the104 participants, leaving the patterns from the majority of participants unidentified. The same issue applied to many

other items in this experiment, as well as many previous studies employing the same method (e.g., Chark et al., 2015; Holcomb & Nelson, 1992; Keren & Roelosfma, 1995; Read, et al., 2016; Sayman & Öncüler, 2009; Scholten & Read, 2006; 2010; Sopher & Sheth, 2006; Weber & Chapman, 2005), which highlighted the importance of testing the (in)dependence between the degree of impatience and the pattern of impatience.

Correlation between DI and impatience. With the quantified degree of impatience and the degree of decreasing impatience at the individual level, Experiment 1 revealed a significant correlation between the overall degree of impatience and degree of decreasing impatience (r = -.33, p < .001), suggesting that impatient individuals were more likely to exhibit decreasing impatience than patient ones (see Figure 4.1b).

This quantified degree of decreasing impatience, however, was subject to a theoretical limitation (Prelec, 2004). The degree of decreasing impatience was approximated by the difference of the numbers of SS choices in no-FED and FED conditions, which was analogous to the difference between the elicited discount rates in the two conditions, However, according to Prelec (2004), the degree of decreasing impatience should be a function of both the decrease of discount rate over time and the discount rate *per se*.[23] To sidestep the bias from our approximation, we reduced the quantified degree of decreasing impatience to three discrete patterns of impatience: decreasing impatience (DI: d < 0), constant impatience (CI: d = 0) and increasing impatience (II: d > 0). A one-way ANOVA was used to compare the means of SS proportions (or degrees of impatience) across the three groups. If the degree of impatience differed across the three groups, it still confirmed a connection between the degree of impatience and the pattern of impatience. The results did suggest the connection between the pattern of impatience and the degree of impatience by showing that the degrees of impatience differed across the three groups, $F(2, 101) = 3.38$, p = .038. A Tukey post-hoc test showed that the SS proportion was higher in the II group

[23] Prelec (2004) states that the degree of decreasing impatience is the "Arrow convexity of the log of the discount function". Based on Prelec's framework, Attema et al. (2010) developed an experimental procedure called Time Trade-Off sequences to measure decreasing impatience. However, Prelec's (2004) framework and Attema et al.'s (2010) implementation are still incompatible with subadditive discounting. Because of the strong evidence for subadditive discounting in the literature (e.g., Kinari et al., 2009; McAlvanah, 2010; Read, 2001; Read & Roelofsma, 2003; Scholten & Read, 2006; 2010), future development of the framework to measure decreasing impatience should take subadditive discounting into consideration.

than in the DI group (p = .031). Other pair-wise comparisons did not approach significance (ps > .25).

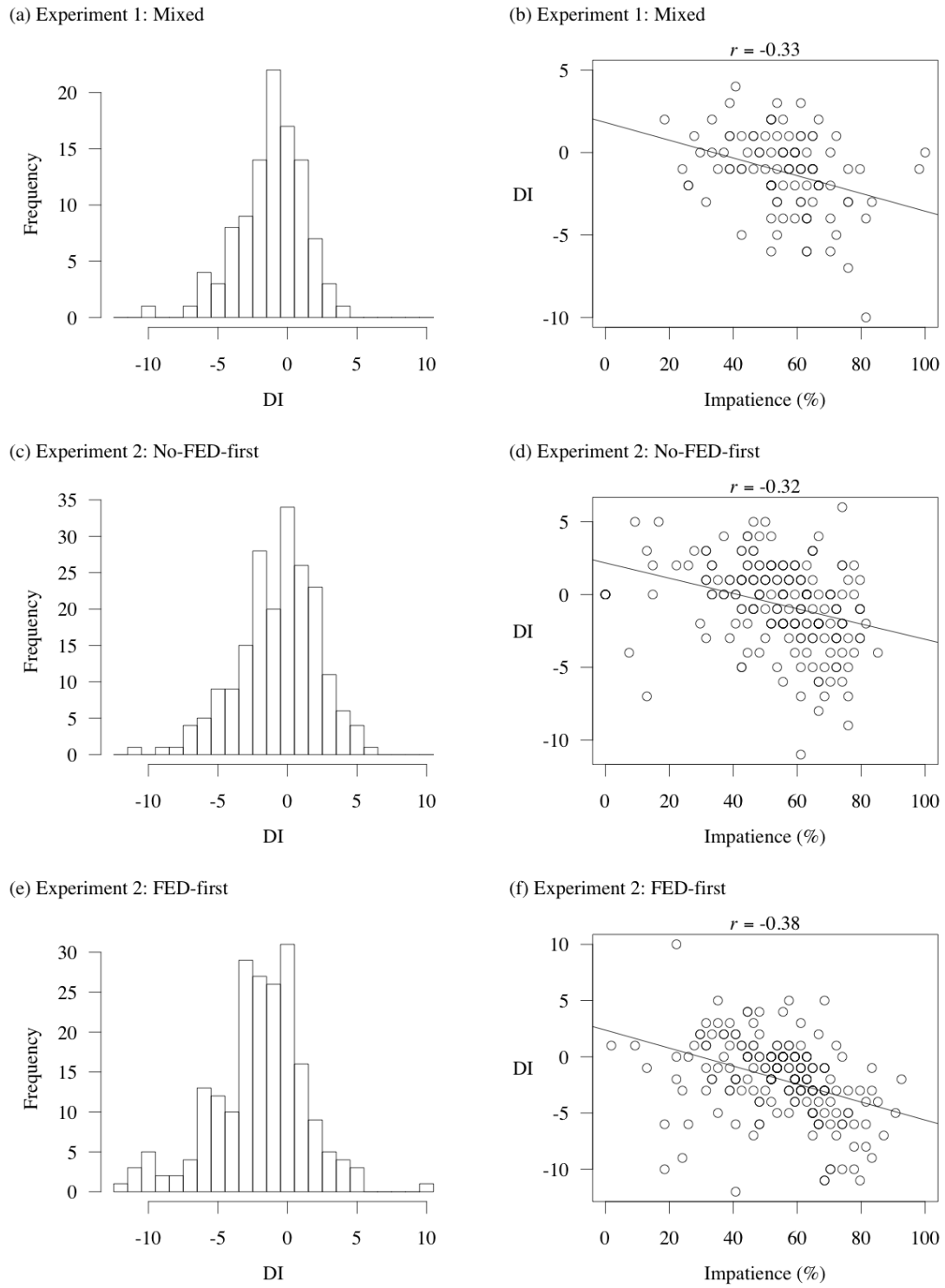

Figure 4.1. Histogram of the degree of decreasing impatience at the individual level (a, c and e) and the correlation between the degree of decreasing impatience and the degree of impatience (b, d and f).

This correlation questioned the assumption that the undetected patterns were homogeneous to the detected patterns with pairs of intertemporal choice items and thus pinpointed the potential design bias in the detected aggregate pattern of impatience with pairs of intertemporal choice items. This was confirmed by the analysis based on pairs of items. As shown in Figure 4.2, the detected patterns of decreasing impatience at the aggregate level clustered among those with unattractive SS options (i.e., the average proportions of SS choices are below 50%). In particular, the average SS proportions from six out of the seven pairs of items that identified decreasing impatience were below 50%.

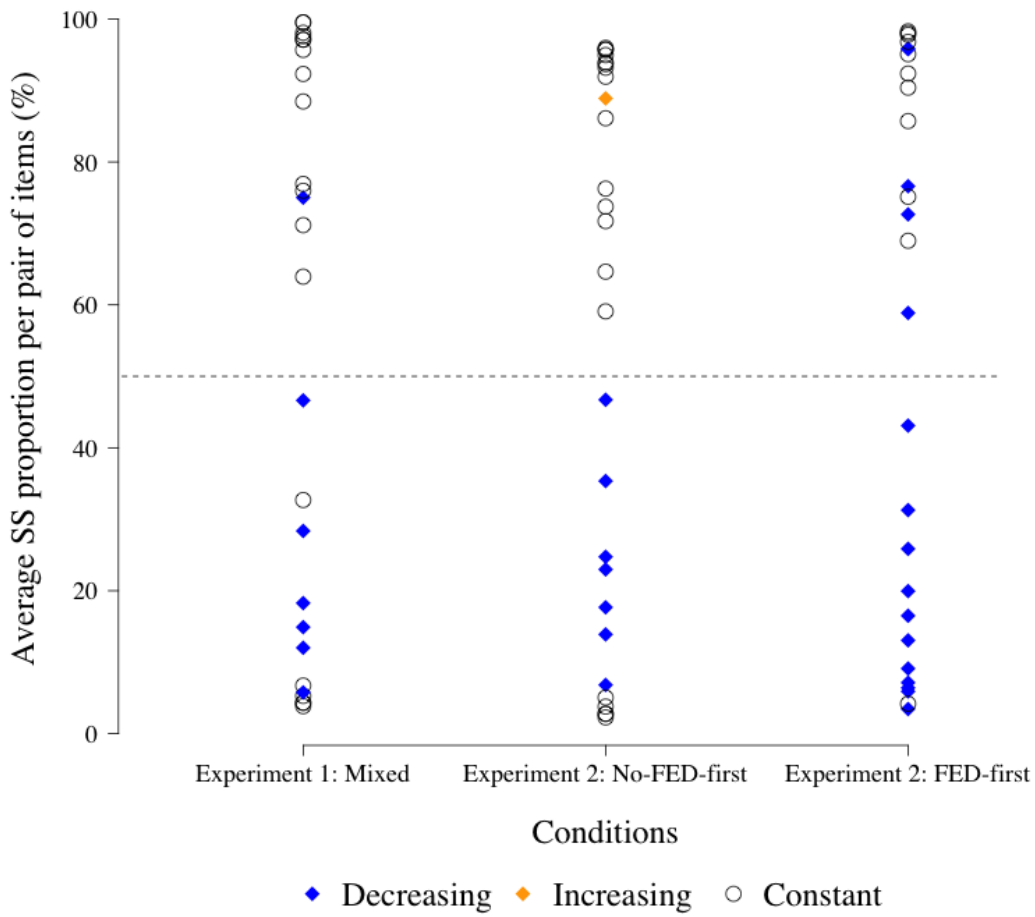

Figure 4.2. Item-based aggregate patterns of impatience. Decreasing impatience or increasing impatience was identified by McNemar's chi-square test with α = .05.

Relationship with other measures. Previous studies showed that cognitive reflection reduced impatience in intertemporal choice (e.g., Frederick, 2005; Chapter

3 in the thesis, see Appendix 3C). To test the relationship between cognitive reflection and impatience in this experiment, participants were divided into two groups according to their performance in the Cognitive Reflection Test (CRT). Those answered all three CRT questions correctly were regarded to have high cognitive reflection and others low cognitive reflection. Consistent with previous findings, participants with high cognitive reflection were less impatient that those with low cognitive reflection, $t(33.3) = 2.99$, $p = 0.006$. The results also revealed a moderate correlation between the impatience measure from intertemporal choice and the self-reported impatience from the Considerations for Future Consequences (CFC) scale, $r = .23$, $p = .017$. The correlation between the degree of decreasing impatience and cognitive reflection or the CFC measure of impatience was not significant (ps > .25).

4.5 Experiment 2: Order Effect

Results from Experiment 1 suggested that the pattern of impatience was not independent from the degree of impatience and that impatient individuals were more likely to be decreasingly impatient that patient individuals. Experiment 2 further examined whether decreasing impatience was stable across decision contexts. The decision context was simply manipulated by varying the sequential order of the no-FED and the FED conditions, so the effect of decision context on the detected pattern of impatience in Experiment 2 was also an order effect.

4.5.1 Experiment 2: Methods

Participants. Four hundred and one adults (232 females, 167 males and others unidentified) from the United Kingdom participated in Experiment 2 via Prolific Academic. Their average age was 32.49 (SD = 11.68) years. Participants were randomly assigned to one of the two between-participant conditions, the no-FED-first and the FED-first conditions, as detailed later. The no-FED-first condition had 198 participants and the FED-first condition had 203 participants.

Design. A 2 (FED: no-FED and FED) × 2 (Order: no-FED-first and FED-first) mixed factorial design was used in Experiment 2. The within-participant factor is the front-end delay (FED) as in Experiment 1, which could be no-FED or FED. The order of the no-FED and the FED conditions were manipulated between participants.

Tasks and procedures. Participants were randomly assigned to one of the two between-participant conditions. As in Experiment 1, participants in both conditions completed three tasks: the intertemporal choice task, the Cognitive Reflection Test (CRT) and the Consideration for Future Consequences (CFC) scale. The intertemporal

choice task included two within-participant conditions, the no-FED condition and the FED condition, each of which consisted of 27 items. So each participant completed 54 intertemporal choice items as in Experiment 1. In the no-FED-first condition, participants answered the no-FED items before the FED items. In the FED-first condition, participants answered the FED-items before the no-FED items. Items in each within-participant condition were presented in a randomized order. After the intertemporal choice task, participants were asked to complete the CRT task and the CFC scale as in Experiment 1. Demographic information was recorded at the end of the experiment (see Appendix 4C for details of the questions).

4.5.2 Experiment 2: Results and discussion

As in Experiment 1, the degree of impatience and the degree of decreasing impatience were quantified in a model-free approach. For each participant, the degree of impatience was quantified by the proportion of SS choices of all 54 intertemporal choice items. The degree of decreasing impatience was quantified by *d*, the number of SS choices in the FED condition minus the number of SS choice in the no-FED condition.

Decreasing impatience. The detected pattern of impatience at the individual level suggested that 47.0% [or 66.0%] of the participants exhibited DI, 35.9% [or 18.7%] exhibited increasing impatience (II) and 17.2% [or 15.3%] exhibited constant impatience (CI) in the no-FED-first [or FED-first] condition (see Figure 4.1c and 4.1e). Aggregately, out of the 27 pairs of intertemporal choices, off-diagonal McNemar's chi-square tests suggested that seven [or fifteen] pairs exhibited decreasing impatience while only one [or no] pair exhibited increasing impatience with the alpha level of 0.05 in the no-FED-first [or FED-first] condition (see Figure 4.2). In summary, Experiment 2 revealed an overall pattern of decreasing impatience at both the individual and the aggregate levels.

Order effect. A multilevel linear regression was used to formally test the effect of the order of presentation:

$$y_i = \alpha_{j[i]} + \beta_0 + \beta_1 x_i^{\text{FED}} + \beta_2 x_i^{\text{Order}} + \beta_3 x_i^{\text{FED}} x_i^{\text{Order}} + \varepsilon_i$$

where the dependent variable, *y*, is the proportion of SS choices in every 27 items per participant (unit: %). x^{FED} is an effect-coded variable indicating the front-end delay (-0.5 = no-FED; 0.5 = FED). x^{Order} was an effect-coded variable indicating the order of

presentation (-0.5 = no-FED-first; 0.5 = FED-first). An individual-level random intercept, $\alpha_{j[i]}$, was used to control for participant-specific variability in her degree of impatience. The model was implemented in *lme4* in R (Bates, Mächler, Bolker, & Walker, 2015). Results are presented under Model (1) in Table 4.4.

The main effect of the front-end delay (FED) confirmed the overall pattern of decreasing impatience (β_1 = -5.00, p < .001). The interaction effect between order and FED suggested the order effect on the degree of decreasing impatience (β_3 = -5.02, p < .001). Participants in the FED-first condition exhibited stronger degrees of decreasing impatience than those in the no-FED-first condition. The order of presentation did not influence the overall degree of impatience, β_2 = 0.88, p > .25.

Table 4.4 Parameter estimation for Experiment 2.

Fixed effect	Parameter	Estimation	
		(1)	(2)
FED	β_1	-5.00***	-4.45***
Order	β_2	0.88	-0.14
FED × Order	β_3	-5.02***	-4.38**
CRT	β_4		-11.36***
FED× CRT	β_5		1.75
Order × CRT	β_6		-2.95
FED × Order × CRT	β_7		2.04
Intercept	β_0	54.53***	51.03***

Note. FED is an effect-coded within-participant variable: -0.5 = the no-FED condition; 0.5 = the FED condition. Order is an effect-coded between-participant variable: -0.5 = the no-FED-first condition; 0.5 = the FED-first condition. CRT is also an effect-coded variable, indicating the level of cognitive reflection: -0.5 = low; 0.5 = high: The dependent variable is the proportion of SS choices of every 27 items per participant (%).

*p < .05; **p < .01; ***p < .001.

Correlation between DI and impatience. The correlation between the degree of decreasing impatience and the degree of impatience was replicated in Experiment 2 in both the no-FED-first and the FED-first conditions. As shown in Figures 4.1d and 4.1f, impatient individuals were more likely to exhibit decreasing impatience than patient ones in both conditions (no-FED-first condition: r = -.32, p < .001; FED-first

condition: r = -.38, p < .001). We further dropped the strength of the quantified decreasing impatience and reduced it to three discrete groups: decreasing impatience (DI: $d < 0$), constant impatience (CI: $d = 0$) and increasing impatience (II: $d > 0$). A 3 (groups: DI, CI and II) × 2 (order: no-FED-first and FED-first) ANOVA suggested that the overall degree of impatience differed across groups ($F(2, 395) = 31.12$, p < .001). Tukey post-hoc tests suggested that participants in the DI group were more impatient than the counterparts in the CI and the II groups (ps < .001). The main effect of presentation order (p > .25) and the interaction effect (p = .067) in the ANOVA were non-significant.

Relationship with other measures. To test if cognitive reflection influenced the degree of impatience and the degree of DI and moderated the order effect, individual participants' level of cognitive reflection (high or low) was involved in a second multilevel linear regression:

$$y_i = \alpha_{j[i]} + \beta_0 + \beta_1 x_i^{\text{FED}} + \beta_2 x_i^{\text{Order}} + \beta_3 x_i^{\text{FED}} x_i^{\text{Order}} + \beta_4 x_i^{\text{CRT}} + \beta_5 x_i^{\text{FED}} x_i^{\text{CRT}} + \beta_6 x_i^{\text{Order}} x_i^{\text{CRT}} + \beta_7 x_i^{\text{FED}} x_i^{\text{Order}} x_i^{\text{CRT}} + \varepsilon_i ,$$

where x_i^{CRT} is the level of cognitive reflection (-0.5 = low cognitive reflection, 0.5 = high cognitive reflection). Results are presented under Model (2) in Table 4.4. Consistent with the findings from Experiment 1, participants with high cognitive reflection were less impatient than those with low cognitive reflection (β_4 = -11.36, p < .001), but cognitive reflection had no effect on the degree of decreasing impatience (β_5 = 1.75, p = .237), nor on the order effect (β_6 = -2.95, p = .477).

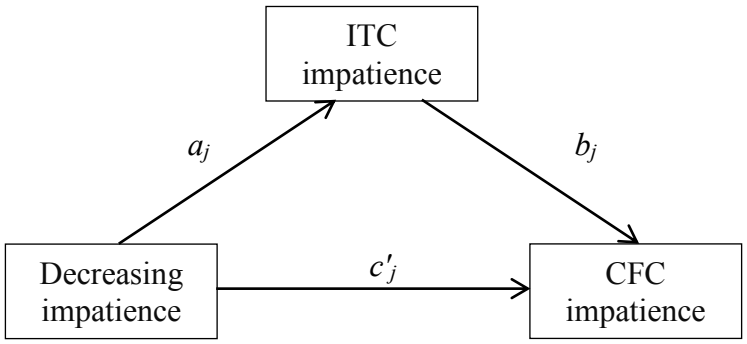

Figure 4.3. Multilevel mediation analysis (Experiment 2). The subscription j refers to the two between-participant conditions. Random effects are allowed for both the direct and the indirect effects in the multilevel mediation model.

The CFC measure of impatience was positively correlated with the impatience measured with the intertemporal choice task in both conditions (no-FED-first condition: r = .33, p < .001; FED-first condition: r = .46, p < .001). It was also negatively correlated with the degree of DI (no-FED-first condition: r = -.16, p = .020; FED-first condition: r = -.23, p < .001). Further multi-level mediation analysis (shown in Figure 4.3; see also Bauer, Preacher, & Gil, 2006) suggested that the correlation between the degree of DI and the CFC measure of impatience could be completely mediated by the degree of impatience measured with the intertemporal choice task (with the mean direct effect $\bar{c}'$ = 0.16, p =.066).

In short, Experiment 2 replicated the correlation between the pattern of impatience and the degree of impatience, which in turn revealed the design bias when using pairs of intertemporal choice items to detect the aggregate pattern of impatience in intertemporal choice. Furthermore, Experiment 2 showed that the degree of decreasing impatience was not a stable trait, but was subject to the decision context, even if the context was simply manipulated by varying the sequential presentation order of intertemporal choice items.

4.6 Discussion

This study focused on the pattern of impatience in intertemporal choice and considered two ways that could influence the detected pattern of impatience. They were a design bias and an order effect. First, the two experiments showed that the individuals' patterns of impatience depend on their degrees of impatience: Impatient individuals were more likely to exhibit decreasing impatience than patient ones. This correlation implied that the use of pairs of intertemporal choice items to detect the aggregate pattern of impatience could be biased because of the design of items (see Figure 4.2 for an overview). Second, the detected pattern of impatience (or the degree of decreasing impatience) was subject to an order effect when the test items were presented in different sequential orders.

4.6.1 Complexity and heterogeneity of the pattern of impatience

A few studies suggested that the pattern of impatience was complex and depended on the length of delays. For example, Sayman and Öncüler (2009) argued that individuals were likely to be increasingly impatient for short delays (e.g., several days) but decreasingly impatient for long delays. Attema et al. (2010) made a similar claim. However, it is important to note that their claims of "short delays" were distinctly different from each other. Attema et al. (2010) considered delays of several

months short, while Sayman and Öncüler (2009) considered delays of the same length long ones. By contrast, Chark et al. (2015) found that people exhibited decreasing impatience for short delays (within approximately one month), but constant impatience for long ones (approximately between one month and one year).Thus, there is not a coherent message in terms of how the pattern changes as a function of the length of delays.

Not only is the pattern of impatience complex, it is also heterogeneous across individuals. However, the importance of this heterogeneity has not been paid enough attention, even among the studies that identified the patterns at the individual level. Particularly, the pattern of impatience was usually studies in isolation, disregarding the degree of impatience. The present study, to our best knowledge, was the first one to test the relationship between individuals' patterns of impatience and their degrees of impatience. We found a connection between the two: Impatient individuals were more likely to exhibit decreasing impatience than patient ones (see Figure 4.1b, d and f).

Importantly, the correlation between the pattern of impatience and the degree of impatience leads to a design bias when pairs of intertemporal choice items between SS and LL were used to detect the aggregate pattern of impatience. Specifically, a pair of items offering attractive LL is more likely to detect decreasing impatience than a pair of items offering attractive SS (see Figure 4.2). This design bias is inevitable when this method is used because it can only identify the patterns from a subset of participants and the design of the pair of items can selectively expose different the patterns from different subsets of participants (e.g., Chark et al., 2015; Holcomb & Nelson, 1992; Keren & Roelosfma, 1995; Read, et al., 2016; Sayman & Öncüler, 2009; Scholten & Read, 2006; 2010; Sopher & Sheth, 2006; Weber & Chapman, 2005). Consequently, studies employing such a method to detect the pattern of impatience are limited in generalisability, especially when only a small number of pairs are involved.

Moreover, the design bias may have interfered the findings on how the pattern of impatience changes over the length of the delay when pairs of intertemporal choice items were used. Although inter-study or intra-study comparisons are not plausible because of numerous confounding factors, such as the magnitude level, front-end delay length and the methodological variations discussed below, future studies may address these confounding factors with systematic experimental designs.

4.6.2 Methodological factors influencing the pattern of impatience

Some studies have shown that some methodological factors can influence the detected pattern of impatience. Read et al. (2005b) found that the framing of intertemporal choice questions influences the detected pattern of impatience. Particularly, decreasing impatience was observed when time was described by the length of delays, but was not when it was described in calendar dates. Studies also suggested that response mode mattered. For example, Read and Roelofsma (2003) found that a matching task was more likely to identify decreasing impatience than a choice task. However, a reversed pattern was observed in two studies using similar item structures: Attema et al. (2010) found an aggregate pattern of increasing impatience with a matching task while Bleichrodt et al. (2016) found an aggregate pattern of decreasing impatience with a choice task in intertemporal choice of monetary gains. As illustrated earlier, Experiment 2 in this study identified a new methodological factor: the order effect on the detected pattern of impatience. Collectively, the effects of the methodological variants suggest that static discount functions with decreasing impatience, increasing impatience or even both do not capture the dynamic feature of the elicited pattern of impatience.

4.6.3 Conclusion, limitations and future directions

This study focused on two methodological concerns on the detected pattern of impatience in intertemporal choice. First, studies using aggregate choices in pairs of intertemporal choice questions only detected a (small) subset of participants and different items might identify the pattern from different subset of participants. The study exhibited an interdependence between the degree of impatience (which is usually termed as time preference or discount rate) and the pattern of impatience (i.e., decreasing, increasing or constant impatience), suggesting that the use of different intertemporal choice items could detect different patterns of impatience. Second, the change of decision context, which was simply manipulated by the sequential presentation order of items, could substantially influence the pattern of impatience. The latter called for a better understanding of the cognitive processes involved in intertemporal choice as well as the pattern of impatience.

One limitation of this study is that we only studied a special case of decreasing impatience, i.e., the immediacy effect, where the base set has no front-end delay (FED). Future investigation could extend this to more general decreasing impatience to compare short-FED and long-FED conditions.

This study also has two recommendations for future research into the elicitation of the pattern of impatience in intertemporal choice. First, when appropriate, researchers should elicit indifference points to identify the pattern of impatience to avoid the potential design bias in using pairs of items (e.g., Attema et al., 2010; Chark et al., 2015). Second, when necessary, instead of using a single pair of item, researchers should use multiples pairs of items that offer extensive relative attractiveness between SS and LL to detect the pattern of impatience from individuals with different degree of impatience.

CHAPTER 5 GENERAL DISCUSSION

The study of intertemporal choice has drawn much attention in both psychology and economics for decades. Undoubtedly there are a vast number of accounts for it (see Frederick et al., 2002; Scholten & Read, 2010; Lempert & Phelps, 2016). Each account speaks a different voice, even if the situation is as simple as a binary choice between a smaller-sooner sum of money and a larger-later sum of money. Among all, one is the "trait" voice, which proposes that time preference is a stable trait that characterise human behaviour in various situations and can be measured with specific intertemporal choice tasks. For example, studies have shown that measured time preference in laboratory experiments can predict human financial behaviour (Ashraf, Karlan, & Yin, 2006), criminal act (Wilson & Daly, 2006), and health and eating behaviour (Appelhans et al., 2011; Borghans & Golsteyn, 2006; Reimers et al., 2009). Another is the "constructive" voice. This voice, by contrast, proposes that intertemporal choice or intertemporal preference is unstable and malleable to diverse framing of intertemporal choice tasks (see *framing effects* in Section 1.2.1 in Chapter 1 [pp. 17]) and others normatively irrelevant factors (e.g., exogenous attention). Although the "constructive" research agenda does not dictate that preferences are exclusively constructed, it pays little attention to the stability of preferences.

The two voices contrast starkly, but very little effort has been made to alleviate the tension. One way to alleviate the tension between the "trait" and the "constructive" voices is to foster a better understanding of the cognitive processes underlying intertemporal choice, therefore introducing models thereof so as to reconcile the conflicting theoretical standing points. Below, I will first summarise the findings from the preceding chapters and then discuss the theoretical implications and potential future directions.

5.1 Summaries of the Findings

In this section, I will briefly recap the findings from the three empirical chapters, crystallising our understanding of the findings.

5.1.1 Evaluation rules

The evaluation rule concerns how information is represented and how choice is made. In Chapter 2, an array of models that embrace different evaluation rules were quantitatively contested based on large-scale secondary data sets. The results from Chapter 2 provide strong support for the class of attribute-based models and thus the attribute-based evaluation rule. The findings are consistent with other studies that compared attribute-based models with alternative-based models in intertemporal choice (e.g., Dai & Busemeyer, 2014; Scholten et al., 2014). Especially, Dai and Busemeyer (2014) compared alternative-based with attribute-based models with individual-level data. They also found that attribute-based models received stronger evidence from a larger number of participants.

5.1.2 Attention effects

Exogenous attention is normatively irrelevant to value-based decision making and should not influence intertemporal choice. However, as shown in Chapter 3 and other studies (e.g., Fisher & Rangel, 2014), exogenous attention influences intertemporal preference in systematic ways, as in other value-based decision making (Armel et al., 2008; Fiedler & Glőckner, 2012; Krajbich et al., 2010; Stewart et al., 2016). Moreover, the attention effects revealed in Chapter 3 are beyond the framework in the literature. While the existing framework usually assumes an option-wise attention effect and sometimes an attribute-wise attention effect, the study in Chapter 3 revealed a component-wise attention effect: Each component receiving attention is overweighed independently from other components in the same option or along the same attribute.

5.1.3 Background contrast effects

Background contrast indicates violations of sequential independence, which means that the preference in one choice should not be influenced by others. Consistent with other studies (e.g., Simonson & Tversky, 1992; Stewart et al., 2015), Chapter 3 reveals robust background contrast effects on intertemporal choice. In addition, the order effect revealed in Chapter 4 is related to the background contrast effect. Collectively, it is evident that background contrast plays an important role in the elicited time preference and decreasing impatience, a stylised behavioural pattern in intertemporal choice. However, this has been largely overlooked.

5.2 Theory Development Revisited

Intertemporal choice modelling started with economic rationality (Fisher, 1930). Later on, it is the behavioural regularities that have motivated the development of the theories and models of intertemporal choice. In this section, I will briefly outline the history of the theory and model development in the study of intertemporal choice, with the implications from the present thesis.

5.2.1 Static models

In the early stage of model development, intertemporal choice models were often built upon specific empirical findings. For example, to accommodate decreasing impatience (or hyperbolic discounting), a class of hyperbolic discounting models were proposed (e.g., Herrnstein, 1981; Mazur, 1987; Laibson, 1997; Loewenstein & Prelec, 1992). To accommodate both decreasing impatience and increasing impatience, more sophisticated discount models have been proposed (Bleichrodt, Rohde, & Wakker, 2009; Ebert & Prelec, 2007; Sayman & Öncüler, 2009). To accommodate the absolute magnitude effect, value functions with increasing elasticity have been proposed (e.g., Chapman, 1996b; Loewenstein & Prelec, 1992; Scholten et al., 2014).

Violations of transitivity in intertemporal choice gave rise to the class of attribute-based models (Leland 2002, Roelofsma & Read, 2000; Rubinstein 2003; Scholten & Read, 2010). Chapter 2 further suggests that the descriptive accuracy of attribute-based models is beyond intransitivity because attribute-based models that cannot accommodate intransitivity in terms of *weak stochastic transitivity* still exhibit much stronger descriptive accuracy than alternative-based ones.[24] This suggests that the attribute-based evaluation rule is probably more psychologically plausible than other evaluation rules (Vlaev, Chater, Stewart, & Brown, 2011). This idea is compatible with a recent eye-tracking study by Arieli et al. (2011). Their process-level eye-tracking data suggest that eye movements during decision making are made more often within attributes than within alternatives in intertemporal choice. This cognitive process strongly suggests that decision makers make attribute-based evaluation more often than alternative-based evaluation in intertemporal choice.

[24] Note that, as shown in Chapter 2, in combination with the Luce specification most attribute-based models can accommodate relative-nonadditivity (an intransitive choice pattern) in terms of the violation of the *product rule*. However, no model, except for the full tradeoff model (TM), predicts intransitivity in terms of the violation of the *weak stochastic transitivity*.

5.2.2 Dynamic models

Recently, dynamic diffusion models have been introduced to account for the stochastic components in value-based decision making including intertemporal choice (e.g., Busemeyer & Townsend, 1993; Usher & McClelland, 2004). For example, Dai and Busemeyer (2014) applied dynamic diffusion models to capture the stochastic variation in the choice process, which is viewed as a preference accumulation process. Correspondingly, a choice is made when the accumulated preference for an option reaches a threshold.

Some other researchers extended dynamic diffusion models with the involvement of the process-level data, though outside of intertemporal choice (e.g., Krajbich et al., 2010; Krajbich & Rangel, 2011; Towal, Mormann, & Koch, 2013). Particularly, Krajbich and colleagues made a crucial assumption that the speed of preference accumulation of an option depends on visual attention. Preference accumulation for an option will be accelerated when it is focused attention on. This approach has embraced great popularity and has accumulated evidence in many studies (Armel et al., 2008; Franco-Watkins et al., 2016; Krajbich et al., 2010; Krajbich & Rangel, 2011; Shimojo et al., 2003; Stewart et al., 2016 Störmer & Alvarez, 2016). Implicitly, these studies held an option-wise attention effect. Following this line of accounts, the study in Chapter 3 further suggests that not only does the attention effect operate in an option-wise way, it also operate in attribute-wise and component-wise ways.

In addition, the study in Chapter 3 reveals background contrast effects on intertemporal choice, symbolising a violation of sequential independence. The order effect found in Experiment 2 of Chapter 4 further suggests a violation of sequential independence. Recently, Lempert, Glimcher and Phelps (2015) and Stewart et al. (2015) also provided evidence on how the revealed discount function, as well as the revealed utility function, for intertemporal choice varies according to the distributions of attribute values (either outcomes or delays) used in the experiment. A few models have offer accounts to the violations of sequential independence. For example, the Decision-by-Sampling model (Stewart, Chater, & Brown, 2006; Stewart et al., 2015; DbS) assumes that the valuation of a quantity is relative but not absolute, which has its deep roots in psychophysics (Laming, 1997; Stewart, Brown, & Chater, 2005). Models such as Decision-by-Sampling is also compatible with a constructive view of preference (Payne, Bettman, & Johnson, 1992).

5.2.3 A theory gap

Models such as Decision-by-Sampling (Stewart et al., 2006) offer theoretical accounts for the constructive features of intertemporal choice or preference. Although they do not dictate that preferences are purely constructed, the inherent stability of preferences is rarely an integral part of these models. By contrast, static intertemporal choice models assume stable intertemporal preference but are unsurprisingly silent on the dynamic features that arise from the elicitation of intertemporal preference. There is a lack of theoretical linkage between the two voices.

A possible solution to filling this gap is the Bayesian approach to human learning (Gopnik &. Tenenbaum, 2007; Oaksford & Chater, 2007; Tenenbaum, Griffiths, & Kemp, 2006). Within the Bayesian approach, time preference can be viewed as a mental construct with uncertainty and can be learnt or updated from repeated choices (Amir & Levav, 2008), interactions with the environment (Rieskamp, Busemeyer, & Mellers, 2006), arbitrary anchoring (Ariely, Loewenstein, & Prelec, 2003) and peer influence (Narayan, Rao, & Saunders, 2011). Accordingly, the decision maker has a *prior* probabilistic distribution of his own time preference. Exposure to a tradeoff between the delays the outcome magnitude is regarded as a signal, or a piece of *evidence*, that decision makers takes into consideration when learning their own preference. With the piece of evidence, the prior distribution can be updated to a *posterior* distribution. This approach may also make a sensible link between the "trait" voice, which focuses on the prior or posterior distributions of the probabilistic time preference, and the "constructive" voice, which focuses on the evidence for preference updating.

5.3 Extensions to Other Domains of Intertemporal Choice

The thesis focuses on binary intertemporal choice between smaller-sooner (SS) and larger-later (LL) monetary gains, such as choosing between $100 in one year or $200 in two years. This is the most intensively employed paradigm in the study of intertemporal choice, but researchers have also used other paradigms to study intertemporal choice. These paradigms vary in different dimensions, including the response mode, single-dated outcomes *vs*. sequences of outcomes, the sign of outcomes (i.e., gains *vs.* losses), commodities (e.g., money *vs.* health). To further understand if the findings from the thesis can be extended to other domains of intertemporal choice, it is essential to consider the differences between these domains and to understand the cognitive processes underpinning those differences. By doing

so, researchers might also be able to judge whether the cognitive processes found to play a role in the binary SS-LL task can speak for the difference between intertemporal choices in different domains.

5.3.1 Other response modes

Although choice tasks are the most widely-used approach, some studies used other response modes such as matching (e.g., Attema et al, 2010; Olivola & Wang, 2016; Read & Roelofsma, 2003) or pricing (e.g., Tversky, Slovic, & Kahneman, 1990) to study intertemporal choice. These studies suggest that response modes play a role in intertemporal choice. In the literature, one of the most prominent theoretical explanations for the role of response modes is *scale compatibility* by Amos Tversky and his colleagues (Slovic, Griffin, & Tversky, 1990; Tversky et al., 1990; Tversky, Sattath, & Slovic, 1988). Scale compatibility means that an attribute looms larger when it is compatible with the response required. For example, when people are asked to indicate the amount of immediately available money equivalent to $100 in one year, the amount attribute is compatible with the response mode and thus its weight looms larger. By contrast, when people are asked to indicate the length of delay to the recipient of $50 that makes it equally valuable to $100 in one year, the delay attribute is compatible with the response mode and then looms larger. Consistent with the theoretical account, Tversky et al. (1990) found that people tended to be less impatient in pricing of delayed amounts, which loomed the amount attribute larger, than in choice between delayed amounts, which was supposed to be a neutral response mode (not triggering either attribute). Similarly, Read and Roelofsma (2003) found that people were less impatient in amount-matching than in binary choice and Olivola and Wang (2016) found that people were less impatient in the amount-matching task than in the delay-matching task.

Scale compatibility is especially relevant to the attention effect in Chapter 3. Indeed, scale compatibility can be re-interpreted as the attribute-wise attention effect as the former suggests that when an attribute is compatible to the response mode, the comparison along that attribute looms and thus is overweighed. I conjecture that this compatibility-led overweight of an attribute could be an effect of contingent attention allocation.

5.3.2 Sequences

Studies also suggest that intertemporal choice of sequences is evaluated differently from intertemporal choice of single-date outcomes (e.g., Chapman, 1996a;

Frederick & Loewenstein, 2008; Gigliotti & Sopher, 1997; Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991; Manzini, Mariotti, & Mittone, 2010; Read & Powell, 2002; Read & Scholten, 2012; Rubinstein, 2003; Scholten, Read, & Sanborn, 2016). The most compelling finding in choice of intertemporal sequences is probably *negative* time preference, which refers to that people prefer a desired outcome to take place later, rather than earlier. For examples, in a choice between sequences in Figures 5.1a and 5.1b. Many people tend to prefer the improving sequence in Figure 5.1a to the deteriorating sequence in Figure 5.1b, although the deteriorating sequence is superior to the improving one given any positive discount rate.

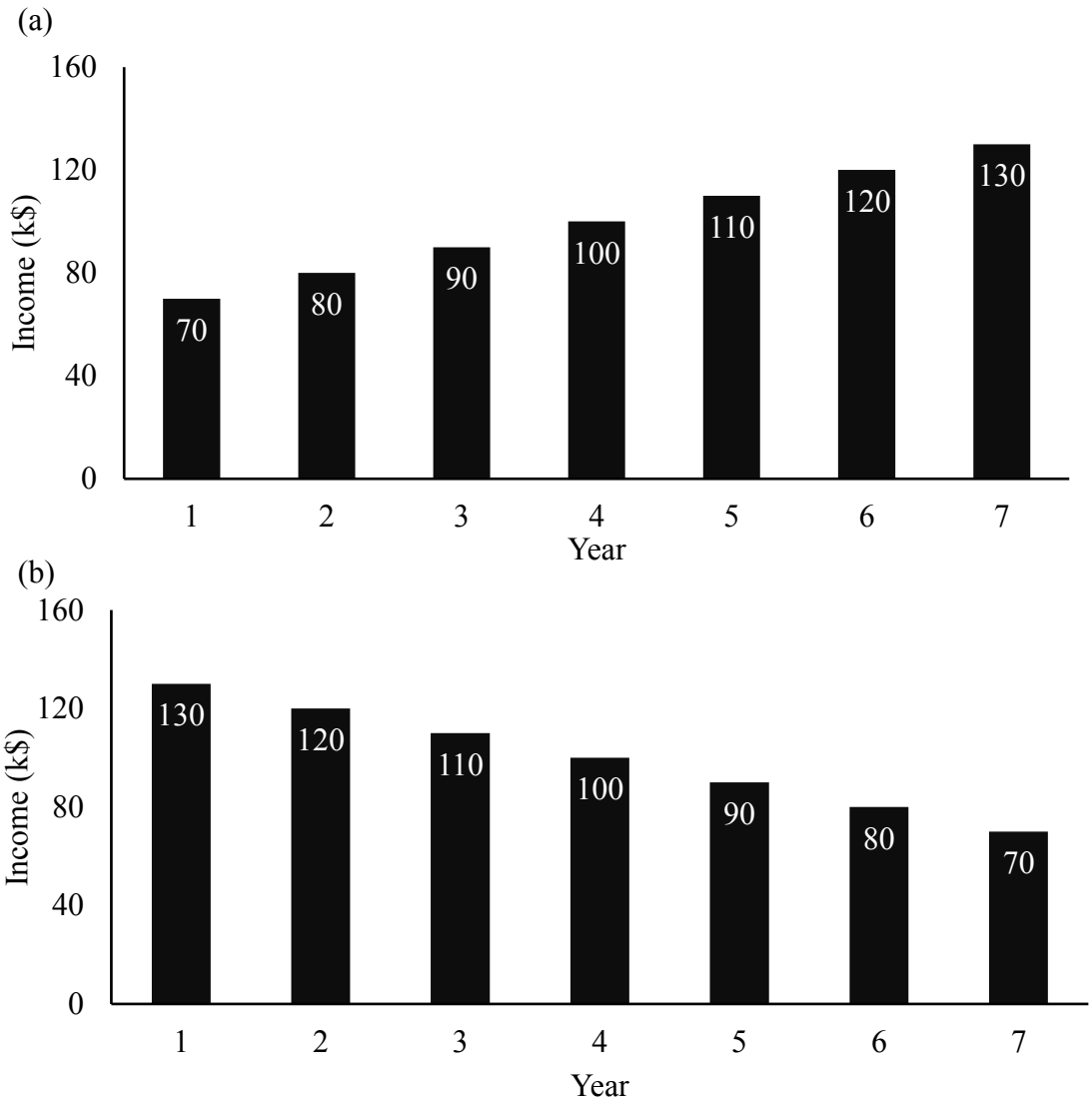

Figure 5.1. Examples of improving and deteriorating sequences with an equal sum: (a) an improving sequence; (b) a deteriorating sequence.

However, negative time preference in intertemporal choice between sequences is not universal either (e.g., Frederick & Loewenstein, 2008; Gigliotti & Sopher, 1997; Manzini, Mariotti & Mittone, 2010; Scholten, Read & Sanborn, 2016). It interacts with many other factors including response mode (Frederick & Loewenstein, 2008). Frederick and Loewenstein (2008) further point out that it is lists of multiple motives that drive the preferences for the improving, the flat and the deteriorating sequences respectively. Specifically, they identified nine factors: (1) anticipation and dread, (2) contrast effects, (3) extrapolation, (4) uncertainty, (5) opportunity cost, (6) pure time preference, (7) diminishing marginal utility, (8) equity among selves and (9) divide equally heuristic. Among them, Factors 1-3 favour improving sequences, 4-6 favour deteriorating sequences and 7-9 favour flat sequences. The resulting preference could be the joint effect by all those motives and should depend on the features of the task that activate these motives.

5.3.3 Losses

Intertemporal choice of monetary payments or losses has also been studied (e.g., Estle et al., 2006; Hardisty et al., 2013; Ohmura et al., 2005; Xu et al., 2009; Yates & Watts, 1975), though much less often than monetary gains. These studies always compared gains with losses and have documented the evidence for the *sign effect* as discussed in Chapter 1. However, the psychological processes underlying intertemporal choice of losses *per se* are rarely discussed. Future research could pay more attention to the potential differences between the psychological processes of intertemporal choice between gains and intertemporal choice between losses.

5.3.4 Non-monetary goods

Intertemporal choice of non-monetary goods is an understudied area despite that studies often suggest commodity of the goods plays a role (Chapman 1996b; Charlton & Fantino, 2008; Estle et al., 2007; Odum et al., 2006; 2011; Odum & Rainaud, 2003; Reuben et al., 2010; Tsukayama & Duckworth, 2010; Ubfal, 2016) and that they are mostly more ecologically valid than intertemporal choice of monetary outcomes. The reasons for the lack of attention to intertemporal choice of non-monetary goods is probably two-fold. First, the illiquidity of non-monetary goods makes them not easily tradable in a market, thus invalidating the Net Present Value (NPV) model to models such decisions. Second, because of satiation, the utility from consumption is not easy to capture for the Discounted Utility (DU) model. As discussed in Chapter 1, the body of research on the discounting of non-monetary goods

always fall prey to such limitations. Future research could pay more attention to understanding the cognitive processes in intertemporal choice of non-monetary goods.

5.4 Conclusion

To conclude, the studies in this thesis cover several aspects of the cognitive processes involved in intertemporal choice, including the evaluation rules, the attention effects and the background contrast effect regarding intertemporal choice in general and decreasing impatience, a specific behavioural pattern in intertemporal choice. The findings not only improve our understanding of the cognitive and psychological underpinnings of this type of choice, but also offer insights to reconcile the conflicting findings documented in the literature.

REFERENCES

Abdellaoui, M., Attema, A. E., & Bleichrodt, H. (2010). Intertemporal tradeoffs for gains and losses: an experimental measurement of discounted utility. *Economic Journal*, *120*(545), 845–866. doi:10.1111/j.1468-0297.2009.02308.x

Ahlbrecht, M. & Weber, M. (1997). An empirical study on intertemporal decision making under risk. *Management Science*, *43* (6), 813–826. doi:10.1016/j.ejor.2006.11.022

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463-496. doi:10.1037/h0076860

Ainslie, G. & Haendel, V. (1983). The motives of the will. *Etiologic Aspects of Alcohol and Drug Abuse* (Charles C. Thomas, Springfield), 119-140.

Ainslie, G., & Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Animal Learning & Behavior*, *9*(4), 476-482. doi:10.3758/BF03209777

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583-618. doi:10.1111/j.1468-0262.2008.00848

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Behavioral econometrics for psychologists. *Journal of Economic Psychology*, *31* (4), 553–576. doi:10.1016/j.joep.2010.03.017

Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, *102*(7), 3333-3356. doi:10.1257/aer.102.7.3333

Appelhans, B. M., Woolf, K., Pagoto, S. L., Schneider, K. L., Whited, M. C., & Liebman, R. (2011). Inhibiting food reward: delay discounting, food reward sensitivity, and palatable food intake in overweight and obese women. *Obesity*, *19*(11), 2175–2182. doi:10.1038/oby.2011.57

Appelt, K. C., Hardisty, D. J., & Weber, E. U. (2011). Asymmetric discounting of gains and losses: A query theory account. *Journal of Risk and Uncertainty*,

43(2), 107-126. doi:10.1007/s11166-011-9125-1

Arieli, A., Ben-Ami, Y., & Rubinstein, A. (2011). Tracking decision makers under uncertainty. *American Economic Journal: Microeconomics*, *3* (4), 68–76. doi:10.1257/mic.3.4.68

Ariely, D., Loewenstein, G., & Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *Quarterly Journal of Economics, 118*(1), 73-105.

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396–403.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *Quarterly Journal of Economics*, *121*(2), 635–672. doi:10.1162/qjec.2006.121.2.635

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. doi:10.1287/mnsc.1100.1219

Augenblick, N., Niederle, M., & Sprenger, C. (2015). Working over time: Dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics*, 1067-1115. doi:10.1093/qje/qjv020

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67* (1), 1–48. doi:10.18637/jss.v067.i01

Baker, F., Johnson, M. W., & Bickel, W. K. (2003). Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *Journal of Abnormal Psychology*, *112* (3), 382–392. doi:10.1037/0021-843X.112.3.382

Baron, J. (2000). Can we use human judgments to determine the discount rate? *Risk Analysis*, *20*(6), 861–868. doi:10.1111/0272-4332.206079

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, *69* (2), 205–223. doi:10.1016/j.geb.2009.11.003

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. *Management Science*, *35*(3), 270–284. doi:10.1287/mnsc.35.3.270

Besag, J. (1989). A candidate's formula: a curious result in Bayesian prediction. *Biometrika*, *76*(1), 183–183.

Bhatia, S. (2014). Sequential sampling and paradoxes of risky choice. *Psychonomic Bulletin & Review, 21*(5), 1095-1111. doi:10.3758/s13423-014-0650-1

Blavatskyy, P. R. & Pogrebna, G. (2010). Models of stochastic choice and decision theories: Why both are important for analyzing decisions. *Journal of Applied Econometrics, 25*(6), 963-986. doi:10.1002/jae.1116

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. doi:10.1007/s11166-016-9240-0

Bleichrodt, H., & Johannesson, M. (2001). Time preference for health: A test of stationarity versus decreasing timing aversion. *Journal of Mathematical Psychology, 45*(2), 265-282. doi:10.1006/jmps.2000.1312

Bleichrodt, H., Rohde, K. I., & Wakker, P. P. (2009). Non-hyperbolic time inconsistency. *Games and Economic Behavior*, 66(1), 27-38. doi:10.1016/j.geb.2008.05.007

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience Theory of Choice Under Risk. *Quarterly Journal of Economics*, *127* (3), 1243–1285. doi:10.1093/qje/qjs018

Borghans, L. & Golsteyn, B. H. H. (2006). Time discounting and the body mass index - evidence from the Netherlands. *Economics and Human Biology*, *4* (1), 39–61. doi:10.1016/j.ehb.2005.10.001

Brealey, R. A., Myers, S. C., Allen, F., & Mohanty, P. (2012). *Principles of Corporate Finance*. New York, NY: McGraw-Hill Education

Brooks, S. P. & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7* (4), 434–455. doi:10.1007/s005850050743

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*(1), 108–132. doi:10.1006/jmps.1999.1279

Busemeyer, J. R. & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. doi:10.1037/0033-295X.100.3.432

Cairns, J. & van der Pol, M. (1997). Saving future lives: A comparison of three discounting models. *Health Economics, 6*(4), 341-350. doi:10.1002/(SICI)1099-1050(199707)6:4<341::AID-HEC277>3.0.CO;2-Y

Cairns, J., & van der Pol, M. (2000). Valuing future private and social benefits: The discounted utility model versus hyperbolic discounting models. *Journal of Economic Psychology, 21*(2), 191-205. doi:10.1016/S0167-4870(99)00042-2

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*(3), 308-313. doi:10.1038/nn1194

Cavagnaro, D. R., Aranovich, G. J., McClure, S. M., Pitt, M. A., & Myung, J. I. (2016). On the functional form of temporal discounting: an optimized adaptive test. *Journal of Risk and Uncertainty*, *52*(3), 233–254. doi:10.1007/s11166-016-9242-y

Cavagnaro, D. R., Regenwetter, M., & Popova, A. (2016, April). *Logits and Fechnerians and Bayes*. Paper presented at the 54th Edwards Bayesian Research Conference, Fullerton, CA.

Chapman, G. B. (1996a). Expectations and preferences for sequences of health and money. *Organizational Behavior and Human Decision Processes*, *67*(1), 59-75. doi:10.1006/obhd.1996.0065

Chapman, G. B. (1996b). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22* (3), 771–791. doi:10.1037/0278-7393.22.3.771

Chark, R., Chew, S. H., & Zhong, S. (2015). Extended present bias: A direct experimental test. *Theory and Decision, 79*(1), 151-165. doi:10.1007/s11238-014-9462-z

Charlton, S. R., & Fantino, E. (2008). Commodity specific rates of temporal discounting: does metabolic function underlie differences in rates of discounting? *Behavioural Processes, 77*(3), 334-342. doi:10.1016/j.beproc.2007.08.002

Cheng, J. & González-Vallejo, C. (2016). Attribute-wise vs. alternative-wise mechanism in intertemporal choice: Testing the proportional difference, trade-off, and hyperbolic models. *Decision, 3*(3), 190-215. doi:10.1037/dec0000046

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62, 73–101. doi:10.1146/annurev.psych.093008.100427

Cohen, A. L., Sanborn, A. N., & Shiffrin, R. M. (2008). Model evaluation using grouped or individual data. *Psychonomic Bulletin & Review, 15*(4), 692-712. doi:10.3758/PBR.15.4.692

Coller, M. & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics*, *2*(2), 107–127. doi:10.1007/BF01673482

Cubitt, R. P., McDonald, R., & Read, D. (2017). Time matters less when outcomes differ: Uni-modal versus cross-modal comparisons in intertemporal choice. *Management Science*. doi:10.1287/mnsc.2016.2613

Cubitt, R. P., & Read, D. (2007). Can intertemporal choice experiments elicit time preferences for consumption? *Experimental Economics*, *10*(4), 369-389. doi:10.1007/s10683-006-9140-2

Dai, J. (2017). Are intertemporal preferences transitive? A Bayesian analysis of repeated individual intertemporal choices. *Decision, 4*(1), 1-24. doi:10.1037/dec0000054

Dai, J. (in press) Using tests of transitivity to evaluate psychological models of intertemporal choice for individuals: Reply to Scholten (2016). *Decision*.

Dai, J. & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, *143*(4), 1489–1514. doi:10.1037/a0035976

Dai, Z., Grace, R. C., & Kemp, S. (2009). Reward contrast in delay and probability discounting. *Learning & Behavior, 37*(3), 281-288. doi:10.3758/lb.37.3.281

Denwood, M. J. (in press). runjags: An R package providing interface utilities, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software*. URL: http://runjags.sourceforge.net

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, *100* (3), 1238–1260. doi:10.1257/aer.100.3.1238

Doyle, J. R. (2013). Survey of time preference, delay discounting models. *Judgment and Decision Making*, *8* (2), 116–135.

Ebert, J. E. (2010). The surprisingly low motivational power of future rewards:

Comparing conventional money-based measures of discounting with motivation-based measures. *Organizational Behavior and Human Decision Processes, 111*(2), 71-92. doi:10.1016/j.obhdp.2009.10.003

Ebert, J. E. & Prelec, D. (2007). The fragility of time: time-insensitivity and valuation of the near and far future. *Management Science*, *53* (9), 1423–1438. doi:10.1287/mnsc.1060.0671

Enzler, H. B., Diekmann, A., & Meyer, R. (2014). Subjective discount rates in the general population and their predictive power for energy saving behavior. *Energy Policy*, *65*, 524–540. doi:10.1016/j.enpol.2013.10.049

Epper, T., Fehr-Duda, H., & Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. J*ournal of Risk and Uncertainty, 43*(3), 169-203. doi:10.1007/s11166-011-9129-x

Ericson, K. M. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting does. *Psychological Science*, *26*(2), 826–833. doi:10.1177/0956797615572232

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science, 18*(1), 58-63. doi:10.1111/j.1467-9280.2007.01849.x

Fiedler, S. & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, *3* (335), 1–18. doi:10.3389/fpsyg.2012.00335

Fishburn, P. C. & Rubinstein, A. (1982). Time preference. *International Economic Review*, *23*(3), 677–694. doi:10.2307/2526382

Fisher, I. (1910) *Elementary Principles of Economics*. NY New York: Macmillan,

Fisher, I. (1930). *The Theory of Interest*. NY New York: Macmillan.

Fisher, G. & Rangel, A. (2014). Intertemporal discount rates are mediated by relative attention. *Unpublished Manuscript*. Retrieved from http://www.anderson.ucla.edu/Documents/areas/fac/marketing/discounting.pdf

Franconeri, S. & Simons, D. (2003). Moving and looming stimuli capture attention. *Perception and Psychophysics*, *65* (7), 999–1010. doi:10.3758/BF03194829

Franco-Watkins, A. M., Mattson, R. E., & Jackson, M. D. (2016). Now or later? Attentional processing and intertemporal choice. *Journal of Behavioral*

Decision Making, *29*(2-3), 206–217. doi:10.1002/bdm.1895

Frank, R. (2014). *Microeconomics and Behaviour* (7th ed. p.139-167). McGraw-Hill Higher Education.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19* (4), 25–42. doi:10.1257/089533005775196732

Frederick, S., & Loewenstein, G. (2008). Conflicting motives in evaluations of sequences. *Journal of Risk and Uncertainty*, *37*(2-3), 221-235. doi:10.1007/s11166-008-9051-z

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40* (2), 351–401. doi:10.1257/002205102320161311

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D.B., Vehtari, A, & Rubin, D. (2014a) Evaluating, comparing, and expanding models. In *Bayesian Data Analysis* (3rd ed., pp. 165-195). Florida, FL: CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D.B., Vehtari, A, & Rubin, D. (2014b) Model checking. In *Bayesian Data Analysis* (3rd ed., pp. 141-164). Florida, FL: CRC Press.

Gelman, A., & Rubin, D. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science, 7*(4), 457-472.

Gigliotti, G., & Sopher, B. (1997). Violations of present-value maximization in income choice. *Theory and Decision*, 43, 45–69. doi:10.1023/A:1004950613488

Gonzalez-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, *109*(1), 137–155. doi:10.1037//0033-295X.109.1.137

Gopnik, A., & Tenenbaum, J. B. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science, 10*(3), 281-287. doi:10.1111/j.1467-7687.2007.00584.x

Green, L., Fristoe, N., & Myerson, J. (1994a). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, *1*(3), 383–389. doi:10.3758/BF03213979

Green, L., Fry, A. F., & Myerson, J. (1994b). Discounting of delayed rewards: A life-span comparison. *Psychological Science*, *5*(1), 33–36. doi:10.1111/j.1467-9280.1994.tb00610.x

Green, L., Myerson, J., & Macaux, E. W. (2005). Temporal discounting when the choice is between two delayed rewards. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 1121. doi:10.1037/0278-7393.31.5.1121

Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica*, *83*(1), 335–352. doi:10.3982/ECTA10872

Han, R., & Takahashi, T. (2012). Psychophysics of time perception and valuation in temporal discounting of gain and loss. *Physica A: Statistical Mechanics and Its Applications*, *391*(24), 6568-6576. doi:10.1016/j.physa.2012.07.012

Hardisty, D. J., Appelt, K. C., & Weber, E. U. (2013a). Good or bad, we want it now: Fixed-cost present for gains and losses explains magnitude asymmetries in intertemporal choice. *Journal Behavioral Decision Making 26*(4), 348-361. doi:10.1002/bdm.1771

Hardisty, D. J., Thompson, K. F., Krantz, D. H., & Weber, E. U. (2013b). How to measure time preferences: An experimental comparison of three methods. *Judgment and Decision Making*, *8*(3), 236–249.

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology: General*, *138*(3), 329-340. doi:10.1037/a0016433

Hare, T. A., Malmaud, J., & Rangel, A. (2011). Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *Journal of Neuroscience*, *31* (30), 11077–11087. doi:10.1523/JNEUROSCI.6383-10.2011

Herrnstein, R. J. (1981). Self-control as response strength. *Quantification of Steady-State Operant Behavior* (Elsevier/North-Holland, Amsterdam), 3-20.

Hey, J. D. (2003). *Intermediate Microeconomics*. New York, NY: McGraw-Hill Education.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental. Economics*. 8(4), 325-345. doi:10.1007/s10683-005-5373-8

Heyman, G. M. (1996). Resolving the contradictions of addiction. *Behavioral and Brain Sciences, 19*(4), 561-574. doi:10.1017/S0140525X00042990

Holcomb, J. H. & Nelson, P. S. (1992). Another experimental look at individual time preference. *Rationality and Society*, *4*(2), 199–220. doi:10.1177/1043463192004002006

Holt, D. D., Green, L., Myerson, J., & Estle, S. J. (2008). Preference reversals with losses. *Psychonomic Bulletin & Review*, *15*(1), 89–95. doi:10.3758/PBR.15.1.89

Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution. In (T. Gilovich, D. Griffin & D. Kahneman, Eds.), *Intuitive Judgment in Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press, pp. 49–81.

Kahneman, D. & Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 363-391. doi::0012-9682(197903)47:2<263:PTAAOD>2.0.CO;2-3

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. Journal of the American Statistical Association, 90(430), 773-795. doi:10.1080/01621459.1995.10476572

Keren, G. & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes*, *63* (3), 287–297. doi:10.1006/obhd.1995.1080

Keynes, J. M (1936). *The General Theory of Employment, Interest and Money*. London: Macmillan (reprinted 2007).

Kinari, Y., Ohtake, F., & Tsutsui, Y. (2009). Time discounting: Declining impatience and interval effect. *Journal of Risk and Uncertainty*, *39* (1), 87–112. doi:10.1007/s11166-009-9073-1

Kirby, K. N. & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, *6* (2), 83–89. doi:10.1111/j.1467-9280.1995.tb00311.x

Kirby, K. N. & Maraković, N. N. (1995). Modelling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes*, *64* (1), 22–30. doi:10.1006/obhd.1995.1086

Kirby, K. N. & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*, *3*(1), 100–104. doi:10.3758/bf03210748

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*(1), 78–87. doi:10.1037/0096-

3445.128.1.78

Kirby, K. N., & Santiesteban, M. (2003). Concave utility, transaction costs, and risk in measuring discounting of delayed rewards. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 66-79.

Knez, M., & Camerer, C. (2000). Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games. *Organizational Behavior and Human Decision Processes, 82*(2), 194-216. doi:10.1006/obhd.2000.2882

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, *28* (2), 287–309. doi:10.2307/1907722

Kőszegi, B. & Szeidl, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics*, *128*(1), 53–104. doi:10.1093/qje/qjs049

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13* (10), 1292–1298. doi:10.1038/nn.2635

Krajbich, I. & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108* (33), 13852–13857. doi:10.1073/pnas.1101328108

Kruschke, J. (2010). *Doing Bayesian Data Analysis: A Tutorial Introduction with R*. Academic Press.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, *112*(2), 443–477. doi:10.1162/003355397555253

Laming, D. R. J. (1997). *The Measurement of Sensation*. Oxford: Oxford University Press.

Laury, S. K., McInnes, M. M., & Swarthout, J. T. (2012). Avoiding the curves: Direct elicitation of time preferences. *Journal of Risk and Uncertainty, 44*(3), 181-217. doi:10.1007/s11166-012-9144-6

LeBoeuf, R. A. (2006). Discount rates for time versus dates: The sensitivity of discounting to time-interval description. *Journal of Marketing Research*, *43*(1), 59-72. doi:10.1509/jmkr.43.1.59

Leland, J. W. (2002). Similarity judgments and anomalies in intertemporal choice. *Economic Inquiry*, *40*(4), 574–581. doi:10.1093/ei/40.4.574

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and

discount rate in intertemporal choice are reference dependent. *Journal of Experimental Psychology: General*, *144*(2), 366–373. doi:10.1037/xge0000047

Lempert, K. M., & Phelps, E. A. (2016). The malleability of intertemporal choice. *Trends in Cognitive Sciences*, *20*(1), 64-74. doi:10.1016/j.tics.2015.09.005

Lerner, J. S., Li, Y., & Weber, E. U. (2012). The financial costs of sadness. *Psychological science, 24*(1): 72-79. doi:10.1177/0956797612450302

Loewenstein, G. (1988). Frames of mind in intertemporal choice. *Management Science*, *34*(2), 200–214. doi:10.1287/mnsc.34.2.200

Loewenstein G (1992) The fall and rise of psychological explanations in the economics of intertemporal choice. In G. Loewenstein & J. Elster (Eds), *Choice Over Time* (pp. 3-34). New York: Russell Sage Foundation.

Loewenstein, G. & Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics*, *107* (2), 573–597. doi:10.2307/2118482.

Loewenstein, G. F., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review*, *100*(1), 91-108. 10.1037/0033-295X.100.1.91

Loewenstein, G., & Sicherman, N. (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics*, 67-84.

Luce R. D. (1959) *Individual Choice Behavior: A Theoretical Analysis*. New York, NY: Wiley.

Madden, G. J., Bickel, W. K., & Jacobs, E. A. (1999). Discounting of delayed rewards in opioid-dependent outpatients: exponential or hyperbolic discounting functions? *Experimental and Clinical Psychopharmacology*, *7*(3), 284–293. doi:10.1037/1064-1297.7.3.284

Magen, E., Dweck, C. S., & Gross, J. J. (2008). The hidden-zero effect: Representing a single choice as an extended sequence reduces impulsive choice. *Psychological Science*, *19*(7), 648–649. doi:10.1111/j.1467-9280.2008.02137.x

Malkoc, S. A., & Zauberman, G. (2006). Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon. *Journal of Marketing Research*, *43*(4), 618-627. doi:10.1509/jmkr.43.4.618

Manzini, P., Mariotti, M., & Mittone, L. (2010). Choosing monetary sequences: Theory and experimental evidence. *Theory and Decision, 69*(3), 327–354. doi:10.1007/s11238-010-9214-7

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *The Effect of Delay and of Intervening Events on Reinforcement Value: Quantitative Analyses of Behavior* (pp. 55–73). Hillsdale, NJ: Lawrence Erlbaum.

McAlvanah, P. (2010). Subadditivity, patience, and utility: The effects of dividing time intervals. *Journal of Economic Behavior & Organization*, *76* (2), 325–337. doi:10.1016/j.jebo.2010.06.009

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *Journal of Neuroscience*, *27* (21), 5796–5804. doi:10.1523/JNEUROSCI.4246-06.2007

McKerchar, T., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., & Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behavioural Processes*, *81*(2), 256–259. doi:10.1016/j.beproc.2008.12.017

Meier, S., & Sprenger, C. D. (2012). Time discounting predicts creditworthiness. *Psychological Science, 23*(1), 56-58. doi:10.1177/0956797611425931

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science, 8*(6), 423-429. doi:10.1111/j.1467-9280.1997.tb00455.x

Michaelson, L., de la Vega, A., Chatham, C., & Munakata, Y. (2013). Delaying gratification depends on social trust. *Frontiers in Psychology*, *4*, 1-7. doi:10.3389/fpsyg.2013.00355

Myerson, J., Baumann, A. A., & Green, L. (2014). Discounting of delayed rewards: (A)theoretical interpretation of the Kirby questionnaire. *Behavioural Processes, 107*, 99-105. doi:10.1016/j.beproc.2014.07.021

Myerson, J. & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64* (3), 263–276. doi:10.1901/jeab.1995.64-263

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior*, *76*(2), 235-243. doi:10.1901/jeab.2001.76-235

Noor, J. (2009). Hyperbolic discounting and the standard model: Eliciting discount functions. *Journal of Economic Theory, 144*(5), 2077-2083. doi:10.1016/j.jet.2009.03.005

Oaksford, M., & Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. UK, Oxford: Oxford University Press.

Odum, A. L. (2011). Delay discounting: trait variable? *Behavioural Processes, 87*(1), 1-9. doi:10.1016/j.beproc.2011.02.007

Odum, A. L., Baumann, A. A., & Rimington, D. D. (2006). Discounting of delayed hypothetical money and food: Effects of amount. *Behavioural Processes, 73*(3), 278-284. doi:10.1016/j.beproc.2006.06.008

Odum, A. L., & Rainaud, C. P. (2003). Discounting of delayed hypothetical money, alcohol, and food. *Behavioural Processes, 64*(3), 305-313. doi:10.1016/S0376-6357(03)00145-1

Ohmura, Y., Takahashi, T., & Kitamura, N. (2005). Discounting delayed and probabilistic monetary gains and losses by smokers of cigarettes. *Psychopharmacology, 182*(4), 508-515. doi 10.1007/s00213-005-0110-8

Olea, J. L. M., & Strzalecki, T. (2014). Axiomatization and measurement of quasi-hyperbolic discounting. *The Quarterly Journal of Economics, 129*(3), 1449-1499. doi:10.1093/qje/qju017

Olivola, C. Y., & Wang, S. W. (2016). Patience auctions: The impact of time vs. money bidding on elicited discount rates. *Experimental Economics, 19*(4), 864–885. doi:10.1007/s10683-015-9472-x

O'Donoghue, T. & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103–124. doi:10.1257/aer.89.1.103

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: a constructive processing perspective. *Annual Review of Psychology*, *43*(1), 87–131. doi:10.1146/annurev.ps.43.020192.000511

Pender, J. L. (1996). Discount rates and credit markets: Theory and evidence from rural India. *Journal of Development Economics, 50*(2), 257-296. doi:10.1016/S0304-3878(96)00400-2

Peters, J., Miedl, S. F., & Büchel, C. (2012). Formal comparison of dual-parameter temporal discounting models in controls and pathological gamblers. *PLoS. ONE*, *7*(11), 1–12. doi:10.1371/journal.pone.0047225

Petry, N. M. (2001). Delay discounting of money and alcohol in actively using

alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology*, *154*(3), 243–250. doi:10.1007/s002130000638

Pine, A., Seymour, B., Roiser, J. P., Bossaerts, P., Friston, K. J., Curran, H. V., & Dolan, R. J. (2009). Encoding of marginal utility across time in the human brain. *Journal of Neuroscience*, *29*(30), 9575–9581. doi:10.1523/JNEUROSCI.1126-09.2009

Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, *113* (1), 57–83. doi:10.1037/0033-295X.113.1.57

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* (Vol. 124, p. 125).

Prelec, D. (2004). Decreasing impatience: A criterion for non-stationary time preference and "hyperbolic" discounting. *Scandinavian Journal of Economics, 106*(3), 511-532. doi:10.1111/j.0347-0520.2004.00375.x

Prelec, D., & Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science, 17*(1), 4-28. doi:10.1287/mksc.17.1.4

Priester, J. R., Dholakia, U. M., & Fleming, M. A. (2004). When and why the background contrast effect emerges: Thought engenders meaning by influencing the perception of applicability. *Journal of Consumer Research, 31*(3), 491-501. doi:10.1086/425084

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org/

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*,111-163. doi:10.2307/271063

Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, *23*(1), 5–32. doi:10.1023/A:1011198414683

Read, D. (2004). Intertemporal choice. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 424–443). New York NY: Wiley.

Read, D., Airoldi, M., & Loewe, G. (2005a). Intertemporal tradeoffs priced in interest rates and amounts: A study of method variance. *Unpublished Manuscript*.

Read, D., Frederick, S., & Airoldi, M. (2012). Four days later in Cincinnati: Longitudinal tests of hyperbolic discounting. *Acta Psychologica*, *140*(2), 177–185. doi:10.1016/j.actpsy.2012.02.010

Read, D., Frederick, S., Orsel, B., & Rahman, J. (2005b). Four score and seven years from now: The date/delay effect in temporal discounting. *Management Science*, *51*(9), 1326–1335. doi:10.1287/mnsc.1050.0412

Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 573–588. doi:10.1037/a0029177.

Read, D., Loewenstein, G., & Kalyanaraman, S. (1999). Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic. *Journal of Behavioral Decision Making, 12*(4), 257. doi:10.1002/(SICI)1099-0771(199912)12:4<257::AID-BDM327>3.0.CO;2-6

Read, D., McDonald, R., & He, L. (in press). Intertemporal choice: Choosing for future. In (A. Lewis Eds.), *The Cambridge Handbook of Psychology and Economic Behaviour* (2nd ed.)

Read, D., Olivola, C. Y., & Hardisty, D. J. (2016). The value of nothing: Asymmetric attention to opportunity costs drives intertemporal decision making. *Management Science*. doi:10.1287/mnsc.2016.2547

Read, D., & Powell, M. (2002). Reasons for sequence preferences. *Journal of Behavioral Decision Making*, *15*(5), 433-460. doi:10.1002/bdm.429

Read, D. & Roelofsma, P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes*, *91*(2), 140–153. doi:10.1016/S0749-5978(03)00060-8

Read, D., & Read, N. L. (2004). Time discounting over the lifespan. *Organizational Behavior and Human Decision Processes, 94*(1), 22-32. doi:10.1016/j.obhdp.2004.01.002

Read, D., & Roelofsma, P. H. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes, 91*(2), 140-153. doi:10.1016/S0749-5978(03)00060-8

Read, D. & Scholten, M. (2012). Tradeoffs between sequences: Weighing accumulated outcomes against outcome-adjusted delays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38* (6), 1675–1688. doi:10.1037/a0028216

Read, D., & van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes, 76*(2), 189-205. doi:10.1006/obhd.1998.2803

Reimers, S., Maylor, E. A., Stewart, N., & Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income, education and real-world impulsive behavior. *Personality and Individual Differences, 47*(8), 973-978. doi:10.1016/j.paid.2009.07.026

Restle, F. (1961). *Psychology of Judgment and Choice: A Theoretical Essay*. New York, NY: John Wiley and Sons.

Reuben, E., Sapienza, P., & Zingales, L. (2010). Time discounting for primary and monetary rewards. *Economics Letters, 106*(2), 125-127. doi:10.1016/j.econlet.2009.10.020

Richards, J. B., Zhang, L., Mitchell, S. H., & de Wit, H. (1999). Delay or probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the Experimental Analysis of Behavior, 71*(2), 121-143. doi:10.1901/jeab.1999.71-121.

Rieskamp, J., Busemeyer, J. R., & Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. *Journal of Economic Literature, 44*(3), 631-661. doi:10.1257/002205106779438006

Rodriguez, M. L., & Logue, A. W. (1988). Adjusting delay to reinforcement: comparing choice in pigeons and humans. *Journal of Experimental Psychology: Animal Behavior Processes, 14*(1), 105. doi:10.1037/0097-7403.14.1.105

Rodriguez, C. A., Turner, B. M., & McClure, S. M. (2014). Intertemporal choice as discounted value accumulation. *PLoS. ONE*, *9* (2), e90138. doi:10.1371/journal.pone.0090138

Roelofsma, P. & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making*, *13*(2), 161–177. doi:10.1002/(SICI)1099-0771(200004/06)13:2<161::AID-BDM348>3.0.CO;2-P

Rubinstein, A. (2003). ‘‘Economics and psychology”? The case of hyperbolic discounting. *International Economic Review*, *44*(4), 1207–1216. doi:10.1111/1468-2354.t01-1-00106

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*(2), 155–161.

Sayman, S. & Onculer, A. (2009). An investigation of time inconsistency. *Management Science*, *55*(3), 470–482. doi:10.1287/mnsc.1080.0942

Scheibehenne, B., Rieskamp, J., & Wagenmakers, E.-J. (2013). Testing adaptive toolbox models: A Bayesian hierarchical approach. *Psychological Review, 120*(1), 39. doi:10.1037/a0030777

Scholten, M. (in press). Thoughts on analyses of intransitive intertemporal preference. *Decision*.

Scholten, M. & Read, D. (2006). Discounting by intervals: A generalized model of intertemporal choice. *Management Science*, *52*(9), 1424–1436. doi:10.1287/mnsc.1060.0534

Scholten, M. & Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological Review*, *117*(3), 925–944. doi:10.1037/a0019619

Scholten, M. & Read, D. (2013). Time and outcome framing in intertemporal tradeoffs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39* (4), 1192–1212. doi:10.1037/a0031171

Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes by time or against time? Evaluation rules in intertemporal choice. *Cognitive Science*, *38*(3), 399–438. doi:10.1111/cogs.12104

Scholten, M., Read, D., & Sanborn, A. (2016). Cumulative weighing of time in intertemporal tradeoffs. *Journal of Experimental Psychology: General*. 145, 1177–1205. doi:10.1037/xge0000198

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. doi:10.1214/aos/1176344136

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*(1), 11-36. doi:10.1016/0010-0277(93)90034-S

Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A., . . . Gray, J. R. (2008). Individual differences in delay discounting relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science*, *19*(9), 904–911. doi:10.1111/j.1467-9280.2008.02175.x

Shelley, M. K. (1993). Outcome signs, question frames and discount rates. *Management Science*, *39*(7), 806-815. doi:10.1287/mnsc.39.7.806

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317-1322. doi:10.1038/nn1150

Simonson, I. & Tversky, A. (1992). Choice in context: tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29* (3), 281–295. doi:10.2307/3172740

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3-22.

Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. *Insights in decision making: A tribute to Hillel J. Einhorn*, 5-27.

Sopher, B., & Sheth, A. (2006). A deeper look at hyperbolic discounting. *Theory and Decision, 60*(2), 219-255. doi:10.1007/s11238-005-4596-7

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London B: Biological Sciences*, *265*(1409), 2015-2020. doi:10.1098/rspb.1998.0534

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639. doi:10.1111/1467-9868.00353

Stevens, J. R. (2016). Intertemporal similarity: Discounting as a last resort. *Journal of Behavioral Decision Making*. doi:10.1002/bdm.1870

Stewart, N., Brown, G. D., & Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1), 3-11. doi:10.1037/0278-7393.28.1.3

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26. doi:10.1016/j.cogpsych.2005.10.003

Stewart, N., Chater, N., Stott, H., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, *132* (1), 23–46. doi:10.1037/0096-3445.132.1.23

Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, 29 (2-3), 116–136. doi:10.1002/bdm.1854

Stewart, N., Reimers, S., & Harris, A. J. L. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to

change their shapes. *Management Science, 61*(3), 687-705. doi:10.1287/mnsc.2013.1853

Störmer, V. S., & Alvarez, G. A. (2016). Attention alters perceived attractiveness. *Psychological Science, 27*(4), 563-571. doi:10.1177/0956797616630964

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty, 32*(2), 101-130. doi:10.1007/s11166-006-8289-6

Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of future consequences: weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66*(4), 742-752. doi:10.1037/0022-3514.66.4.742

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies, 23*(3), 165-180.

Takahashi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Medical Hypotheses, 65*(4), 691-693. doi:10.1016/j.mehy.2005.04.040

Takahashi, T., Oono, H., & Radford, M. H. (2008). Psychophysics of time perception and intertemporal choice models. *Physica A: Statistical Mechanics and its Applications*, *387*(8), 2066–2074. doi:10.1016/j.physa.2007.11.047

Takeuchi, K. (2011). Non-parametric test of time consistency: Present bias and future bias. *Games and Economic Behavior, 71*(2), 456-478. doi:10.1016/j.geb.2010.05.005

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*(7), 309-318. doi:10.1016/j.tics.2006.05.009

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, *8*(3), 201–207. doi:10.1016/0165-1765(81)90067-7

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics, 112*(2), 647-661. doi:10.1162/003355397555226

Toubia, O., Johnson, E., Evgeniou, T., & Delquié, P. (2013). Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters. *Management Science, 59*(3), 613-640. doi:10.1287/mnsc.1120.1570

Towal, R. B., Mormann, M., & Koch, C. (2013). Simultaneous modelling of visual saliency and value computation improves predictions of economic choice. *Proceedings of the National Academy of Sciences*, *110* (40), 3858–3867. doi:10.1073/pnas.1304429110

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Science of the United States of America*, *109* (24), 9659–9664. doi:10.1073/pnas.1119569109

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological review, 117*(4), 1275-1291. doi:10.1037/a0020580

Tsukayama, E., & Duckworth, A. L. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making, 5*(2), 72-82.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, *76* (1), 31–48. doi:10.1037/h0026750

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299. doi:10.1037/h0032955

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*(3), 371–384. doi:10.1037//0033-295X.95.3.371

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *American Economic Review, 8*(1), 204-217.

Ubfal, D. (2016). How general are time preferences? Eliciting good-specific discount rates. *Journal of Development Economics, 118*, 150-170. doi:10.1016/j.jdeveco.2015.07.007

Ungemach, C., Stewart, N., & Reimers, S. (2011). How incidental values from the environment affect decisions about money, risk, and delay. *Psychological Science, 22*(2), 253-260. doi:10.1177/0956797610396225

Usher, M. & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111* (3), 757–769. doi:10.1037/0033-295X.111.3.757

Vlaev, I., & Chater, N. (2006). Game relativity: How context influences strategic decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 131-149. doi:10.1037/0278-7393.32.1.131

Vlaev, I., & Chater, N. (2007). Context effects in games: Local versus global sequential effects on choice in the prisoner's dilemma game. *Judgment and Decision Making, 2*(6), 380.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15* (11), 546–554. doi:10.1016/j.tics.2011.09.008

Vlaev, I., Chater, N., & Stewart, N. (2007). Financial prospect relativity: Context effects in financial decision-making under risk. *Journal of Behavioral Decision Making*, *20* (3), 273–304. doi:10.1002/bdm.555

Vlaev, I., Seymour, B., Dolan, R. J., & Chater, N. (2009). The price of pain and the value of suffering. *Psychological Science*, *20* (3), 309–317. doi:10.1111/j.1467-9280.2009.02304.x

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General, 144*(1), 7. doi:10.1037/xge0000039

Weber, B. J., & Chapman, G. B. (2005). The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect? *Organizational Behavior and Human Decision Processes, 96*(2), 104-118. doi:0.1016/j.obhdp.2005.01.001

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, 60, 53- 85. doi:10.1146/annurev.psych.60.110707.163633

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: a query-theory account. *Psychological Science*, *18*(6), 516–523. doi:10.1111/j.1467-9280.2007.01932.x

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. *Research in Experimental Economics, 12*, 197-292.

Wills, A. J., Reimers, S., Stewart, N., Suret, M., & McLaren, I. P. L. (2000). Tests of the ratio rule in categorization. *The Quarterly Journal of Experimental Psychology: Section A, 53*(4), 983-1011. doi:10.1080/713755935

Wilson, M. & Daly, M. (2006). Are juvenile offenders extreme future discounters? *Psychological Science*, *17*(11), 989–994. doi:10.1111/j.1467-

9280.2006.01817.x

Wu, C. Y., & He, G. B. (2012). The effects of time perspective and salience of possible monetary losses on intertemporal choice. *Social Behavior and Personality: An International Journal, 40*(10), 1645-1653. doi:10.2224/sbp.2012.40.10.1645

Xu, L., Liang, Z. Y., Wang, K., Li, S., & Jiang, T. (2009). Neural mechanism of intertemporal choice: from discounting future gains to future losses. *Brain Research*, *1261*, 65-74. doi:10.1016/j.brainres.2008.12.061

Yates, J. F., & Watts, R. A. (1975). Preferences for deferred losses. *Organizational Behavior and Human Performance, 13*(2), 294-306. doi:10.1016/0030-5073(75)90051-3

Yi, R., Landes, R. D., & Bickel, W. K. (2009). Novel models of intertemporal valuation: Past and future outcomes. *Journal of Neuroscience, Psychology, & Economics.* 2(2), 102-111. doi:10.1037/a0017571

Zauberman, G., Kim, B. K., Malkoc, S. A., & Bettman, J. R. (2009). Discounting time and time discounting: subjective time perception and intertemporal preferences. *Journal of Marketing Research*, *46* (4), 543–556. doi:10.1509/jmkr.46.4.543

APPENDICES

Appendix 2A

Method for maximum likelihood search

Maximum log-likelihood, $\ln(L_{qrs})$, was searched for each combination of model q, stochastic specification r and data set s. In principle, the search was implemented by the SIMPLEX algorithm with the *fminsearch* function in Matlab. However, because of both the high dimensionality of some models and the complexity and heterogeneity in the data sets, three more steps were taken to ensure the identification of the global maximum log-likelihood:

1. The *fminsearch* function was repeated 20 times consecutively to guarantee the search for local optimum;
2. Step 1 was run for 100 independent repetitions;
3. Twenty of the 100 repetitions were preceded by the Metropolis-Hastings version of the Markov Chain Monte Carlo (MCMC) sampler for 50 thousand iterations each, which started with a random sample of parameters.

The search domains were as indicated in Table 2.2. For parameters having an unbounded domain (i.e., $+\infty$), $+\infty$ was replaced by 10^{10} as the search domain.

A total of 11,776 (256 data sets crossed by 46 models) maximum log-likelihood values were obtained with this procedure. The 46 models included 15 core models crossed by 3 stochastic specifications plus 1 baseline model.

Appendix 2B

Summary statistics of the 256 data sets involved in the model comparison

Table A1 below shows the full list of the data sets involved in the model comparison in Chapter 2. The list consists of 256 data sets from 97 research articles or projects. Column *Eligible* indicates if a data set was eligible and thus included in the formal results. A *Yes* indicates that the data set passed the screening procedure based on the relative performance of formal models and the baseline model. A *No* indicates that the data set was excluded from formal analysis because the baseline model outperformed the formal models for the data set. Column *Source* shows the author(s) and the year of publication of the article or research project. Note that a research paper or project always appears more than once in the list because they mostly involved multiple experiments and/or multiple experimental conditions and different experiments/conditions were coded into different data sets. Column *Experiment* indicates the experiment ID from a research article or project. When only one experiment was involved, it is coded as 1. Column *Number of Items*, as the name suggests, indicates the number of binary intertemporal choice items in a data set. Column *Number of Participants* indicates the average number of participants' responses across intertemporal choice items in a data set. Column *Now-Later* indicates whether a data set involved intertemporal choice(s) between an immediately available option and a delayed option. Column *Later-Later* indicates whether a data set involved intertemporal choice(s) between two delayed options.

Table A1 List of the data sets involved in the model comparison in Chapter 2.

Eligible	Source	Experiment	Number of Items	Number of Participants	Now-Later	Later-Later
Yes	Acheson et al. (2011)	1	27	292.96	Yes	No
Yes	Acheson et al. (2011)	1	27	202.93	Yes	No
Yes	Andersen et al. (2008)	1	60	241.97	No	Yes
Yes	Andreoni and Sprenger (2012)	1	21	93.00	Yes	Yes
Yes	Appelhans et al. (2011)	1	147	95.88	Yes	No
Yes	Ashraf et al. (2006)	1	4	1764.00	Yes	Yes
Yes	Augustine and Larsen (2011)	1	27	33.96	Yes	No
Yes	Augustine and Larsen (2011)	1	27	36.00	Yes	No
Yes	Augustine and Larsen (2011)	2	27	30.96	Yes	No
Yes	Augustine and Larsen (2011)	2	27	36.00	Yes	No
Yes	Bauer et al. (2012)	1	10	544.00	No	Yes
Yes	Benjamin et al. (2010a)	1	46	35.93	Yes	Yes
Yes	Benjamin et al. (2010a)	1	46	38.00	Yes	Yes
Yes	Benjamin et al. (2010a)	2	24	54.96	Yes	Yes
Yes	Benjamin et al. (2010a)	2	24	166.00	Yes	Yes
Yes	Benjamin et al. (2010a)	2	24	95.00	Yes	Yes
Yes	Benjamin et al. (2010a)	2	24	72.92	Yes	Yes
Yes	Benjamin et al. (2010a)	2	24	119.92	Yes	Yes
Yes	Benjamin et al. (2010b)	1	24	211.00	Yes	Yes
Yes	Benjamin et al. (2010b)	1	24	210.00	Yes	Yes
Yes	Benjamin et al. (2013)	1	5	92.00	Yes	No
Yes	Benjamin et al. (2013)	2	10	103.00	Yes	Yes
Yes	Benjamin et al. (2013)	3	10	36.90	Yes	Yes
Yes	Benningfield et al. (2014)	1	27	19.00	Yes	No

Yes	Bjork et al. (2004)	1	105	41.00	Yes	No
Yes	Bjork et al. (2004)	1	105	31.00	Yes	No
Yes	Bjork et al. (2004)	1	105	88.00	Yes	No
Yes	Borghans and Golsteyn (2006)	1	10	1150.00	Yes	Yes
Yes	Brase and Brase (2012)	1	27	255.00	Yes	No
Yes	Campitelli and Labollita (2010)	1	3	157.00	Yes	No
Yes	Chao et al. (2009)	1	27	184.59	Yes	No
Yes	Chark et al. (2015)	1	20	125.00	No	Yes
Yes	Chark et al. (2015)	3	30	64.00	No	Yes
Yes	Chark et al. (2015)	2	20	1138.00	No	Yes
Yes	Chen (2010 [Chinese])	1a	2	57.50	Yes	Yes
Yes	Cherniawsky and Holroyd (2013)	1	308	36.00	Yes	No
Yes	Cherniawsky and Holroyd (2013)	1	308	36.00	Yes	No
Yes	Cheung (2014)	1	40	81.00	No	Yes
Yes	Chuang and Schechter (2015)	1	2	449.00	Yes	No
Yes	Chuang and Schechter (2015)	2	2	195.00	Yes	No
Yes	de Wit et al. (2007)	1	126	830.00	Yes	No
Yes	Dittrich and Leipold (2014)	1	6	1019.00	No	Yes
Yes	Dohmen et al. (2010)	1	20	500.00	Yes	No
Yes	Doyle and Chen (2012)	1	33	64.00	Yes	No
Yes	Doyle and Chen (2012)	2	33	38.00	Yes	No
Yes	Doyle and Chen (2012)	3	40	37.00	Yes	No
Yes	Drichoutis and Nayga Jr. (2013a)	1	30	86.00	No	Yes
Yes	Drichoutis and Nayga Jr. (2013b)	1	30	33.00	No	Yes
Yes	Drichoutis and Nayga Jr. (2013b)	1	30	23.00	No	Yes
Yes	Drichoutis and Nayga Jr. (2013b)	1	30	30.00	No	Yes
Yes	Duquette et al. (2012)	1	3	69.33	Yes	No

Yes	Eckel et al. (2007)	1	5	885.00	No	Yes
Yes	Enzler et al. (2014)	1	10	3157.00	Yes	No
Yes	Epper et al. (2011)	1	38	112.00	No	Yes
Yes	Ericson et al. (2015)	1	718	6.53	Yes	Yes
Yes	Ericson et al. (2015)	1	720	6.82	Yes	Yes
Yes	Ericson et al. (2015)	1	719	6.76	Yes	Yes
Yes	Ericson et al. (2015)	1	720	6.48	Yes	Yes
Yes	Ericson et al. (2015)	1	718	5.58	Yes	Yes
Yes	Faralla et al. (2012)	1	120	25.00	Yes	Yes
Yes	Field et al. (2006)	1	182	29.86	Yes	No
Yes	Field et al. (2006)	1	182	90.00	Yes	No
Yes	Fowler and Kam (2006)	1	20	245.00	No	Yes
Yes	Freeman et al. (2013)	1	182	16.00	Yes	No
Yes	Freeman et al. (2013)	1	182	16.00	Yes	No
Yes	Fuchs (1982)	1	6	504.00	Yes	No
Yes	Griskevicius et al. (2012)	2	20	82.00	No	Yes
Yes	Griskevicius et al. (2012)	2	20	61.00	No	Yes
Yes	Griskevicius et al. (2012)	2	20	62.00	No	Yes
Yes	Han and Takahashi (2012)	1	273	50.00	Yes	No
Yes	Han and Takahashi (2012)	1	273	50.00	Yes	No
Yes	Hardisty and Weber (2009)	1	8	64.00	Yes	No
Yes	Hardisty and Weber (2009)	2	8	116.00	Yes	No
Yes	Hardisty et al. (2013)	1	229	1.48	Yes	No
Yes	Hardisty et al. (2013)	1	27	116.00	Yes	No
Yes	Hardisty et al. (2013)	2	3	87.00	Yes	No
Yes	Hardisty et al. (2013)	2	24	98.00	Yes	No
Yes	Hepler et al. (2012)	1	27	29.00	Yes	No

Yes	Holcomb and Nelson (1992)	1	36	98.00	Yes	Yes
Yes	Janis and Nock (2009)	2	27	44.00	Yes	No
Yes	Jimura et al. (2011)	1	6	20.00	Yes	No
Yes	Jimura et al. (2011)	1	6	19.00	Yes	No
Yes	Jimura et al. (2011)	2	6	20.00	Yes	No
Yes	Jimura et al. (2011)	2	6	21.00	Yes	No
Yes	Joireman et al. (2005)	2	27	106.85	Yes	No
Yes	Kassam et al. (2008)	1	8	34.00	Yes	No
Yes	Keren and Roelofsma (1995)	1	2	60.00	Yes	Yes
Yes	Keren and Roelofsma (1995)	3	2	75.00	No	Yes
Yes	Kinari et al. (2009)	1	180	209.00	Yes	Yes
Yes	Kirby and Finch (2010)	1	9	404.67	Yes	No
Yes	Kirby and Marakovic (1996)	1	21	621.00	Yes	No
Yes	Kirby and Petry (2004)	1	27	69.00	Yes	No
Yes	Kirby and Petry (2004)	1	27	44.00	Yes	No
Yes	Kirby and Petry (2004)	1	27	65.00	Yes	No
Yes	Kirby et al. (1999)	1	27	60.00	Yes	No
Yes	Kirby et al. (1999)	1	27	56.00	Yes	No
Yes	Kirby et al. (2002)	1	8	510.25	Yes	No
Yes	Koff and Lucas (2011)	1	21	192.00	No	Yes
Yes	Kosse and Pfeiffer (2012)	1	40	213.00	Yes	No
Yes	Lange and Eggert (2014)	1	176	70.00	Yes	No
Yes	Lerner et al. (2013)	1	27	33.00	Yes	No
Yes	Lerner et al. (2013)	1	27	86.00	Yes	No
Yes	Lerner et al. (2013)	1	27	83.00	Yes	No
Yes	Lerner et al. (2013)	3	41	99.00	Yes	Yes
Yes	Lerner et al. (2013)	3	41	104.00	Yes	Yes

Yes	Lerner et al. (2013)	2	11	62.00	Yes	No
Yes	Lerner et al. (2013)	2	11	68.00	Yes	No
Yes	Lerner et al. (2013)	2	11	60.00	Yes	No
Yes	Li (2008)	1	8	25.00	No	Yes
Yes	Li (2008)	1	8	30.00	No	Yes
Yes	Li (2008)	1	8	25.00	No	Yes
Yes	Li (2008)	1	8	25.00	No	Yes
Yes	Li et al. (2010)	1	2	100.00	Yes	Yes
Yes	Li et al. (2011)	1	3	101.67	Yes	No
Yes	Li et al. (2011)	2	3	103.00	Yes	No
Yes	Litvin and Brandon (2010)	1	27	174.81	Yes	No
Yes	Lucas and Koff (2010)	1	21	184.00	No	Yes
Yes	Lukoseviciute (2011)	1	38	22.50	Yes	Yes
Yes	Lukoseviciute (2011)	1	38	26.00	Yes	Yes
Yes	Lukoseviciute (2011)	1	38	22.00	Yes	Yes
Yes	Lukoseviciute (2011)	1	38	24.00	Yes	Yes
Yes	Manzini et al. (2014)	1	39	64.00	No	Yes
Yes	Manzini et al. (2014)	1	39	64.00	No	Yes
Yes	Manzini et al. (2014)	1	39	64.00	No	Yes
Yes	McAlvanah (2010)	1	20	55.00	Yes	Yes
Yes	McKerchar et al. (2013)	1	336	55.00	Yes	No
Yes	Michaelson et al. (2013)	1	21	78.00	Yes	No
Yes	Michaelson et al. (2013)	1	20	78.00	Yes	No
Yes	Michaelson et al. (2013)	1	20	85.80	Yes	No
Yes	Michaelson et al. (2013)	2	48	43.00	Yes	No
Yes	Michaelson et al. (2013)	2	49	46.00	Yes	No
Yes	Michaelson et al. (2013)	2	49	49.00	Yes	No

Yes	Pepper and Nettle (2013)	1	20	598.00	Yes	No
Yes	Petry et al. (2002)	1	27	64.00	Yes	No
Yes	Petry et al. (2002)	1	27	58.00	Yes	No
Yes	Pulcu et al. (2014)	1	13	29.00	Yes	No
Yes	Pulcu et al. (2014)	1	13	27.85	Yes	No
Yes	Pulcu et al. (2014)	1	13	25.85	Yes	No
Yes	Ramos et al. (2013)	1	9	67.00	No	Yes
Yes	Ramos et al. (2013)	1	9	64.00	No	Yes
Yes	Read et al. (2005)	1	4	28.00	No	Yes
Yes	Read et al. (2005)	3	4	30.00	No	Yes
Yes	Sayman and Onculer (2009)	1	12	38.00	Yes	Yes
Yes	Sayman and Onculer (2009)	2a	16	72.00	Yes	Yes
Yes	Sayman and Onculer (2009)	2b	16	58.50	Yes	Yes
Yes	Sayman and Onculer (2009)	3	18	73.00	Yes	Yes
Yes	Scholten and Read (2010)	1	8	196.00	No	Yes
Yes	Scholten et al. (2014)	2	22	412.36	No	Yes
Yes	Scholten et al. (2014)	3	27	518.00	Yes	No
Yes	Senecal et al. (2012)	1	50	92.00	Yes	No
Yes	Senecal et al. (2012)	2	102	19.96	Yes	No
Yes	Senecal et al. (2012)	2	102	19.98	Yes	No
Yes	Senecal et al. (2012)	2	102	15.98	Yes	No
Yes	Senecal et al. (2012)	2	102	19.97	Yes	No
Yes	Senecal et al. (2012)	2	102	19.99	Yes	No
Yes	Senecal et al. (2012)	2	102	17.94	Yes	No
Yes	Senecal et al. (2012)	2	102	19.92	Yes	No
Yes	Senecal et al. (2012)	2	102	19.98	Yes	No
Yes	Senecal et al. (2012)	2	102	16.00	Yes	No

Yes	Senecal et al. (2012)	2	102	18.79	Yes	No
Yes	Senecal et al. (2012)	2	102	18.44	Yes	No
Yes	Senecal et al. (2012)	2	102	18.95	Yes	No
Yes	Senecal et al. (2012)	3	51	19.90	Yes	No
Yes	Senecal et al. (2012)	3	51	19.94	Yes	No
Yes	Senecal et al. (2012)	3	99	8.73	Yes	No
Yes	Senecal et al. (2012)	4	51	15.96	Yes	No
Yes	Senecal et al. (2012)	4	51	16.00	Yes	No
Yes	Senecal et al. (2012)	4	51	15.96	Yes	No
Yes	Senecal et al. (2012)	4	51	16.00	Yes	No
Yes	Senecal et al. (2012)	4	51	15.84	Yes	No
Yes	Senecal et al. (2012)	4	101	7.99	Yes	No
Yes	Senecal et al. (2012)	4	51	15.90	Yes	No
Yes	Senecal et al. (2012)	4	51	15.98	Yes	No
Yes	Shead and Hodgins (2009)	1	6	60.00	Yes	No
Yes	Stephens and Krupka (2006)	1	4	3509.25	Yes	No
Yes	Stewart et al. (2015)	3a	120	19.00	No	Yes
Yes	Stewart et al. (2015)	3a	120	21.00	No	Yes
Yes	Stewart et al. (2015)	3b	150	9.00	No	Yes
Yes	Stewart et al. (2015)	3b	150	10.00	No	Yes
Yes	Stewart et al. (2015)	4	120	15.00	No	Yes
Yes	Stewart et al. (2015)	4	120	15.00	No	Yes
Yes	Stillwell and Tunney (2012)	1	98	8.53	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	5.09	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	9.82	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	187.17	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	221.36	Yes	No

Yes	Stillwell and Tunney (2012)	1	98	186.76	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	275.39	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	275.06	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	284.62	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	428.74	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	414.99	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	389.76	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	25.72	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	24.98	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	25.43	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	52.28	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	81.58	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	71.31	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	84.32	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	98.08	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	113.21	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	106.89	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	91.68	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	93.06	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	2998.44	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	2990.30	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	2958.17	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	28.87	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	26.13	Yes	No
Yes	Stillwell and Tunney (2012)	1	98	28.53	Yes	No
Yes	Sutter et al. (2013)	2	152	659.75	Yes	Yes
Yes	Torres et al. (2013)	1	27	21.00	Yes	No

Yes	Torres et al. (2013)	1	27	24.00	Yes	No
Yes	Torres et al. (2013)	1	27	20.00	Yes	No
Yes	van der Wal et al. (2013)	1	7	23.00	Yes	No
Yes	van der Wal et al. (2013)	1	7	24.00	Yes	No
Yes	van der Wal et al. (2013)	2	27	15.33	Yes	No
Yes	van der Wal et al. (2013)	3	7	21.00	Yes	No
Yes	van der Wal et al. (2013)	3	7	22.00	Yes	No
Yes	Woebert and Riedl (2013)	1	32	229.47	Yes	Yes
Yes	Xu et al. (2009)	1	40	20.00	Yes	Yes
Yes	Yamane et al. (2013)	1	75	17.00	Yes	Yes
Yes	Yamane et al. (2013)	1	75	50.00	Yes	Yes
Yes	Yamane et al. (2013)	1	75	50.00	Yes	Yes
No	Benjamin et al. (2010a)	1	46	37.00	Yes	Yes
No	Benjamin et al. (2010a)	1	46	46.96	Yes	Yes
No	Chen (2010 [Chinese])	2	3	50.00	No	Yes
No	Chen (2010 [Chinese])	2	4	50.00	No	Yes
No	Chen (2010, Chinese)	2	4	50.00	No	Yes
No	Chen and He (2012, Chinese)	1	2	58.00	Yes	Yes
No	Dai and Fishbach (2013)	1	2	28.00	No	Yes
No	Dai and Fishbach (2013)	2a	2	48.00	No	Yes
No	Enzler et al. (2014)	2	2	1339.50	Yes	Yes
No	Enzler et al. (2014)	3	2	212.50	Yes	Yes
No	Enzler et al. (2014)	3	2	749.50	Yes	Yes
No	Espín et al. (2012)	1	20	160.00	No	Yes
No	Franco-Watkins et al. (2010)	1	80	53.00	Yes	No
No	Franco-Watkins et al. (2010)	1	80	53.00	Yes	No
No	Li (2008)	1	8	24.00	No	Yes

No	Poulos (2010)	1	20	94.00	Yes	No
No	Read et al. (2005)	1	4	30.00	No	Yes
No	Read et al. (2005)	1	4	28.00	No	Yes
No	Read et al. (2005)	3	4	29.00	No	Yes
No	Read et al. (2005)	4	4	28.00	No	Yes
No	Read et al. (2005)	4	4	30.00	No	Yes
No	Read et al. (2005)	4	4	30.00	No	Yes
No	Read et al. (2012)	1	10	128.00	Yes	Yes
No	Read et al. (2012)	2	2	76.00	Yes	Yes
No	Read et al. (2012)	2	2	65.00	Yes	Yes
No	Read et al. (2012)	2	2	60.00	Yes	Yes
No	Rubenstein (2003)	1	2	228.00	No	Yes
No	Scholten and Read (2010)	2	3	233.00	No	Yes
No	van der Wal et al. (2013)	2	27	15.33	Yes	No
No	van der Wal et al. (2013)	2	27	14.00	Yes	No
No	Weber and Champan (2005)	1	2	111.00	Yes	Yes

References

Acheson, A., Vincent, A. S., Sorocco, K. H., & Lovallo, W. R. (2011). Greater discounting of delayed rewards in young adults with family histories of alcohol and drug use disorders: studies from the Oklahoma family health patterns project. *Alcoholism: Clinical and Experimental Research, 35*(9), 1607-1613. doi:10.1111/j.1530-0277.2011.01507.x

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583-618. doi:10.1111/j.1468-0262.2008.00848

Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, *102*(7), 3333-3356. doi:10.1257/aer.102.7.3333

Appelhans, B. M., Woolf, K., Pagoto, S. L., Schneider, K. L., Whited, M. C., & Liebman, R. (2011). Inhibiting food reward: delay discounting, food reward sensitivity, and palatable food intake in overweight and obese women. *Obesity*, *19*(11), 2175–2182. doi:10.1038/oby.2011.57

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics*, *121*(2), 635–672. doi:10.1162/qjec.2006.121.2.635

Augustine, A. A., & Larsen, R. J. (2011). Affect regulation and temporal discounting: Interactions between primed, state, and trait affect. *Emotion, 11*(2), 403-412. doi:10.1037/a0021777

Bauer, M., Chytilová, J., & Morduch, J. (2012). Behavioral foundations of microcredit: Experimental and survey evidence from rural India. *American Economic Review, 102*(2), 1118-1139. doi:10.1257/aer.102.2.1118

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010a). Social identity and preferences (No. w13309). *National Bureau of Economic Research*.

Benjamin, D. J., Choi, J. J., & Fisher, G. W. (2010b). Religious identity and economic behavior? (No. w15925). *National Bureau of Economic Research*.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association, 11*(6), 1231-1255. doi:10.1111/jeea.12055

Benningfield, M. M., Blackford, J. U., Ellsworth, M. E., Samanez-Larkin, G. R., Martin, P. R., Cowan, R. L., & Zald, D. H. (2014). Caudate responses to

reward anticipation associated with delay discounting behavior in healthy youth. *Developmental Cognitive Neuroscience, 7*, 43-52. doi:10.1016/j.dcn.2013.10.009

Bjork, J. M., Hommer, D. W., Grant, S. J., & Danube, C. (2004). Impulsivity in abstinent alcohol-dependent patients: relation to control subjects and type 1–/type 2–like traits. *Alcohol, 34*(2), 133-150. doi:10.1016/j.alcohol.2004.06.012

Borghans, L. & Golsteyn, B. H. H. (2006). Time discounting and the body mass index - evidence from the Netherlands. *Economics and Human Biology*, *4* (1), 39–61. doi:10.1016/j.ehb.2005.10.001

Brase, G. L., & Brase, S. L. (2012). Emotional regulation of fertility decision making: What is the nature and structure of "baby fever"? *Emotion, 12*(5), 1141-1154. doi:10.1037/a0024954

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making, 5*, 182-191.

Chao, L. W., Szrek, H., Pereira, N. S., & Pauly, M. V. (2009). Time preference and its relationship with age, health, and survival probability. *Judgment and Decision Making, 4(*1), 1-19.

Chark, R., Chew, S. H., & Zhong, S. (2015). Extended present bias: A direct experimental test. *Theory and Decision, 79*(1), 151-165. doi:10.1007/s11238-014-9462-z

Chen, H. (2012 [Chinese]). The effect of psychological distance on intertemporal choice (Doctoral dissertation, Zhejiang University, China)

Chen, H., & He, G., (2012 [Chinese]). Psychological mechanism underlying preference reversal in intertemporal choice. Journal of Psychological Science, 35(4),862-867

Chen, Y. (2010 [Chinese]). *The effects of delay description on intertemporal choice under different elicitation modes and frames* (Master's thesis, East China Normal University, China).

Cherniawsky, A. S., & Holroyd, C. B. (2013). High temporal discounters overvalue immediate rewards rather than undervalue future rewards: An event-related brain potential study. *Cognitive, Affective, & Behavioral Neuroscience,13*(1), 36-45. doi:10.3758/s13415-012-0122-x

Cheung, S. L. (2015). Comment on "Risk preferences are not time preferences": On the elicitation of time preference under conditions of risk. *American Economic Review, 105*(7), 2242-2260. doi:10.1257/aer.20120946

Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. Jo*urnal of Development Economics, 117*, 151-170. doi:10.1016/j.jdeveco.2015.07.008

Dai, X., & Fishbach, A. (2013). When waiting to choose increases patience. *Organizational Behavior and Human Decision Processes, 121*(2), 256-266. doi:10.1016/j.obhdp.2013.01.007

de Wit, H., Flory, J. D., Acheson, A., McCloskey, M., & Manuck, S. B. (2007). IQ and nonplanning impulsivity are independently associated with delay discounting in middle-aged adults. *Personality and Individual Differences, 42*(1), 111-121. doi:10.1016/j.paid.2006.06.026

Dittrich, M., & Leipold, K. (2014). Gender differences in time preferences. *Economics Letters, 122*(3), 413-415. doi:10.1016/j.econlet.2014.01.002

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review, 100*(3), 1238-1260. doi:10.1257/aer.100.3.1238

Doyle, J., & Chen, C. (2012). The wages of waiting and simple models of delay discounting. Available at SSRN 2008283.

Drichoutis, A. C., & Nayga Jr, R. M. (2013a). A reconciliation of time preference elicitation methods. Retrieved from https://mpra.ub.uni-muenchen.de/46916/

Drichoutis, A. C., & Nayga Jr, R. M. (2013b). Eliciting risk and time preferences under induced mood states. *Journal of Socio-Economics, 45*, 18-27. doi:10.1016/j.socec.2013.04.008

Duquette, E., Higgins, N., & Horowitz, J. (2012). Farmer discount rates: experimental evidence. *American Journal of Agricultural Economics, 94*(2), 451-456. doi:10.1093/ajae/aar067

Eckel, C. C., Johnson, C., Montmarquette, C., & Rojas, C. (2007). Debt aversion and the demand for loans for postsecondary education. *Public Finance Review, 35*(2), 233-262. doi:10.1177/1091142106292774

Enzler, H. B., Diekmann, A., & Meyer, R. (2014). Subjective discount rates in the general population and their predictive power for energy saving behavior.

Energy Policy, *65*, 524–540. doi:10.1016/j.enpol.2013.10.049

Epper, T., Fehr-Duda, H., & Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. J*ournal of Risk and Uncertainty, 43*(3), 169-203. doi:10.1007/s11166-011-9129-x

Ericson, K. M. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? simple heuristics explain intertemporal choices better than delay discounting does. *Psychological Science*, *26*(2), 826–833. doi:10.1177/0956797615572232

Espín, A. M., Brañas-Garza, P., Herrmann, B., & Gamella, J. F. (2012). Patient and impatient punishers of free-riders. *Proceedings of the Royal Society B: Biological Sciences, 279*(1749), 4923-4928. doi:10.1098/rspb.2012.2043

Faralla, V., Benuzzi, F., Nichelli, P., & Dimitri, N. (2012). Gains and losses in intertemporal preferences: A behavioural study. In A. Innocenti & A. Sirigu (Eds.), *Neuroscience and the Economics of Decision Making* (pp. 146-161). London: Routledge.

Field, M., Santarcangelo, M., Sumnall, H., Goudie, A., & Cole, J. (2006). Delay discounting and the behavioural economics of cigarette purchases in smokers: the effects of nicotine deprivation. *Psychopharmacology, 186*(2), 255-263. doi:10.1007/s00213-006-0385-4

Field, M., Christiansen, P., Cole, J., & Goudie, A. (2007). Delay discounting and the alcohol Stroop in heavy drinking adolescents. *Addiction, 102*(4), 579-586. doi:10.1111/j.1360-0443.2007.01743.x

Fowler, J. H., & Kam, C. D. (2006). Patience as a political virtue: Delayed gratification and turnout. Political Behavior, 28(2), 113-128. doi:10.1007/s11109-006-9004-7

Franco-Watkins, A. M., Rickard, T. C., & Pashler, H. (2010). Taxing executive processes does not necessarily increase impulsive decision making. *Experimental Psychology, 57*(3), 193-201. doi:10.1027/1618-3169/a000024

Freeman, T. P., Morgan, C. J., Brandner, B., Almahdi, B., & Curran, H. V. (2013). Dopaminergic involvement in effort-based but not impulsive reward processing in smokers. *Drug and Alcohol Dependence, 130*(1), 109-114. doi:10.1016/j.drugalcdep.2012.10.016

Fuchs, V. R. (1982). Time preference and health: An exploratory study. In V. R. Fuchs (Eds.), *Economic Aspects of Health* (pp. 93-120). Chicago, IL: U.

Chicago Press.

Griskevicius, V., Tybur, J. M., Ackerman, J. M., Delton, A. W., Robertson, T. E., & White, A. E. (2012). The financial consequences of too many men: Sex ratio effects on saving, borrowing, and spending. *Journal of Personality and Social Psychology, 102*(1), 69-80. doi:10.1037/a0024761

Han, R., & Takahashi, T. (2012). Psychophysics of time perception and valuation in temporal discounting of gain and loss. *Physica A: Statistical Mechanics and Its Applications*, *391*(24), 6568-6576. doi:10.1016/j.physa.2012.07.012

Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: money versus the environment. *Journal of Experimental Psychology: General*, *138*(3), 329-340. doi:10.1037/a0016433

Hardisty, D. J., Thompson, K. F., Krantz, D. H., & Weber, E. U. (2013). How to measure time preferences: An experimental comparison of three methods. *Judgment and Decision Making*, *8*(3), 236–249.

Hepler, J., Albarracin, D., McCulloch, K. C., & Noguchi, K. (2012). Being active and impulsive: The role of goals for action and inaction in self-control. *Motivation and Emotion, 36*(4), 416-424. doi:10.1007/s11031-011-9263-4

Holcomb, J. H. & Nelson, P. S. (1992). Another experimental look at individual time preference. *Rationality and Society*, *4*(2), 199–220. doi:10.1177/1043463192004002006

Janis, I. B., & Nock, M. K. (2009). Are self-injurers impulsive?: Results from two behavioral laboratory studies. Psychiatry Research, 169(3), 261-267. doi:10.1016/j.psychres.2008.06.041

Jimura, K., Myerson, J., Hilgard, J., Keighley, J., Braver, T. S., & Green, L. (2011). Domain independence and stability in young and older adults' discounting of delayed rewards. *Behavioural Processes, 87*(3), 253-259. doi:10.1016/j.beproc.2011.04.006

Joireman, J., Sprott, D. E., & Spangenberg, E. R. (2005). Fiscal responsibility and the consideration of future consequences. *Personality and Individual Differences, 39*(6), 1159-1168. doi:10.1016/j.paid.2005.05.002

Kassam, K. S., Gilbert, D. T., Boston, A., & Wilson, T. D. (2008). Future anhedonia and time discounting. *Journal of Experimental Social Psychology, 44*(6), 1533-1537. doi:10.1016/j.jesp.2008.07.008

Keren, G. & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice.

Organizational Behavior and Human Decision Processes, *63* (3), 287–297. doi:10.1006/obhd.1995.1080

Kinari, Y., Ohtake, F., & Tsutsui, Y. (2009). Time discounting: declining impatience and interval effect. *Journal of Risk and Uncertainty*, *39*(1), 87–112. doi:10.1007/s11166-009-9073-1

Kirby, K. N., & Finch, J. C. (2010). The hierarchical structure of self-reported impulsivity. *Personality and Individual Differences, 48*(6), 704-713. doi:10.1016/j.paid.2010.01.019

Kirby, K. N. & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic Bulletin & Review*, *3*(1), 100–104. doi:10.3758/bf03210748

Kirby, K. N., & Petry, N. M. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction, 99*(4), 461-471. doi:10.1111/j.1360-0443.2003.00669.x

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*, *128*(1), 78–87. doi:10.1037/0096-3445.128.1.78

Kirby, K. N., Godoy, R., Reyes-García, V., Byron, E., Apaza, L., Leonard, W., ... & Wilkie, D. (2002). Correlates of delay-discount rates: Evidence from Tsimane' Amerindians of the Bolivian rain forest. *Journal of Economic Psychology, 23*(3), 291-316. doi:10.1016/S0167-4870(02)00078-8

Koff, E., & Lucas, M. M. (2011). Mood moderates the relationships between impulsiveness and delay discounting. *Personality and Individual Differences, 50*, 1018-1022. doi:10.1016/j.paid.2011.01.016

Kosse, F., & Pfeiffer, F. (2012). Impatience among preschool children and their mothers. *Economics Letters, 115*(3), 493-495. doi:10.1016/j.econlet.2011.12.115

Lange, F., Eggert, F. (2014). Sweet delusion. Glucose drinks fail to counteract ego depletion. *Appetite, 75*, 54-63. doi:10.1016/j.appet.2013.12.020

Lerner, J. S., Li, Y., & Weber, E. U. (2012). The financial costs of sadness. *Psychological Science, 24*(1): 72-79. doi:10.1177/0956797612450302

Li, X. (2008). The effects of appetitive stimuli on out-of-domain consumption impatience. *Journal of Consumer Research, 34*(5), 649-656.

doi:10.1086/521900

Li, S., Su, Y., & Sun, Y. (2010). The effect of pseudo-immediacy on intertemporal choices. *Journal of Risk Research, 13*(6), 781-787. doi:10.1080/13669870903551704

LI, J. Z., Li, S., & Liu, H. (2011). How has the Wenchuan earthquake influenced people's intertemporal choices? *Journal of Applied Social Psychology, 41*(11), 2739-2752. doi:10.1111/j.1559-1816.2011.00847.x

Litvin, E. B., & Brandon, T. H. (2010). Testing the influence of external and internal cues on smoking motivation using a community sample. *Experimental and Clinical Psychopharmacology, 18*(1), 61-70. doi:10.1037/a0017414

Lucas, M. M., & Koff, E. (2010). Delay discounting is associated with 2D:4D ratio in women but not men. *Personality and Individual Differences, 48*, 182-186. doi:10.1016/j.paid.2009.10.002

Lukoseviciute, R. (2011). *The effect of background mood and pre-existing risk on risk and time preferences* (Master's thesis, Tilburg University, the Netherlands). Retrieved from https://www.netspar.nl/assets/uploads/042_RM_Rasa_Lukoseviciute.pdf

Manzini, P., Mariotti,M. and Mittone, L. (2014). A case of framing effects: the elicitation of time preferences. Unpublished Manuscript. Retrieved from http://www.st-andrews.ac.uk/economics/repecfiles/4/1405.pdf

McAlvanah, P. (2010). Subadditivity, patience, and utility: the effects of dividing time intervals. *Journal of Economic Behavior & Organization, 76*(2), 325–337. doi:10.1016/j.jebo.2010.06.009

McKerchar, T. L., Pickford, S., & Robertson, S. E. (2013). Hyperboloid discounting of delayed outcomes: Magnitude effects and the gain--loss asymmetry. *Psychological Record, 63*(3), 441-451.

Michaelson, L., de la Vega, A., Chatham, C., & Munakata, Y. (2013). Delaying gratification depends on social trust. *Frontiers in Psychology*, *4*, 1-7. doi:10.3389/fpsyg.2013.00355

Pepper, G. V., & Nettle, D. (2013). Death and the time of your life: experiences of close bereavement are associated with steeper financial future discounting and earlier reproduction. *Evolution and Human Behavior, 34*(6), 433-439. doi:10.1016/j.evolhumbehav.2013.08.004

Petry, N. M., Kirby, K. N., & Kranzler, H. R. (2002). Effects of gender and family

history of alcohol dependence on a behavioral task of impulsivity in healthy subjects. *Journal of Studies on Alcohol and Drugs, 63*(1), 83-90.

Poulos, J. (2010, January 5). Intertemporal choice. *Experimental Turk: A Blog on Social Science Experiments on Amazon Mechanical Turk*. Retrieved from https://experimentalturk.wordpress.com/2010/01/05/intertemporal-choice/

Pulcu, E., Trotter, P. D., Thomas, E. J., McFarquhar, M., Juhasz, G., Sahakian, B. J., ... & Elliott, R. (2014). Temporal discounting in major depressive disorder. *Psychological Medicine, 44*(9), 1825-1834. doi:10.1017/S0033291713002584

Ramos, D., Victor, T., Seidl-de-Moura, M. L., & Daly, M. (2013). Future discounting by slum-dwelling youth versus university students in Rio de Janeiro. *Journal of Research on Adolescence, 23*(1), 95-102. doi:10.1111/j.1532-7795.2012.00796.x

Read, D., Frederick, S., & Airoldi, M. (2012). Four days later in Cincinnati: Longitudinal tests of hyperbolic discounting. *Acta Psychologica*, *140*(2), 177–185. doi:10.1016/j.actpsy.2012.02.010

Read, D., Frederick, S., Orsel, B., & Rahman, J. (2005). Four score and seven years from now: the date/delay effect in temporal discounting. *Management Science*, *51*(9), 1326–1335. doi:10.1287/mnsc.1050.0412

Rubinstein, A. (2003). ‘‘Economics and psychology”? The case of hyperbolic discounting. *International Economic Review*, *44*(4), 1207–1216. doi:10.1111/1468-2354.t01-1-00106

Sayman, S. & Onculer, A. (2009). An investigation of time inconsistency. *Management Science*, *55*(3), 470–482. doi:10.1287/mnsc.1080.0942

Scholten, M. & Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological Review*, *117*(3), 925–944. doi:10.1037/a0019619

Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes by time or against time? Evaluation rules in intertemporal choice. *Cognitive Science*, *38*(3), 399–438. doi:10.1111/cogs.12104

Senecal, N., Wang, T., Thompson, E., & Kable, J. W. (2012). Normative arguments from experts and peers reduce delay discounting. *Judgment and Decision Making, 7*(5), 568-589.

Shead, N. W., & Hodgins, D. C. (2009). Probability discounting of gains and losses: Implications for risk attitudes and impulsivity. *Journal of the Experimental*

Analysis of Behavior, 92(1), 1-16. doi:10.1901/jeab.2009.92-1

Stephens Jr, M., & Krupka, E. L. (2006). Subjective discount rates and household behavior. *Unpublished Manuscript*. Retrieved from https://www.researchgate.net/profile/Erin_Krupka/publication/228831803_Subjective_Discount_Rates_and_Household_Behavior/links/546ccd890cf2193b94c57710.pdf

Stewart, N., Reimers, S., & Harris, A. J. L. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science, 61*(3), 687-705. doi:10.1287/mnsc.2013.1853

Stillwell, D. J., & Tunney, R. J. (2012). Effects of measurement methods on the relationship between smoking and delay reward discounting. *Addiction, 107*(5), 1003-1012. doi:10.1111/j.1360-0443.2011.03742.x [and also Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist, 70*(6), 543-556. doi:10.1037/a0039210]

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), 510-531. doi:10.1257/aer.103.1.510

Torres, A., Catena, A., Megías, A., Maldonado, A., Cándido, A., Verdejo-García, A., & Perales, J. C. (2013). Emotional and non-emotional pathways to impulsive behavior and addiction. *Frontiers in Human Neuroscience, 7*(43), 1-11. doi:10.3389/fnhum.2013.00043

van der Wal, A. J., Schade, H. M., Krabbendam, L., & van Vugt, M. (2013). Do natural landscapes reduce future discounting in humans? *Proceedings of the Royal Society B: Biological Sciences, 280*(1773), 1-6. doi:10.1098/rspb.2013.2295

Wölbert, E. M & Riedl, A. M., (2013). Measuring Time and Risk Preferences: Reliability, Stability, Domain Specificity. *Unpublished Manuscript*. Retrieved from https://www.econstor.eu/bitstream/10419/79662/1/cesifo_wp4339.pdf

Xu, L., Liang, Z. Y., Wang, K., Li, S., & Jiang, T. (2009). Neural mechanism of intertemporal choice: from discounting future gains to future losses. *Brain Research*, *1261*, 65-74. doi:10.1016/j.brainres.2008.12.061

Yamane, S., Yoneda, H., Takahashi, T., Kamijo, Y., Komori, Y., Hiruma, F., & Tsutsui, Y. (2013). Smokers, smoking deprivation, and time discounting. *Journal of Socio-Economics, 45*, 47-56. doi:10.1007/978-4-431-55402-8_10

Weber, B. J., & Chapman, G. B. (2005). The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect? *Organizational Behavior and Human Decision Processes, 96*(2), 104-118. doi:0.1016/j.obhdp.2005.01.001

Appendix 2C

Parameter Space Partitioning Search

Some intertemporal choice models have irregular forms, so it is difficult to analytically derive their properties in accommodating the behavioural regularities. In such cases, Parameter Space Partitioning (Pitt et al., 2006), implemented by Markov chains, is applied to search the models' parameter spaces, as well as stimulus spaces, to identify if a model is able to accommodate a specific behavioural regularity.

Double-exponential discounting model. The qualitative properties of the double-exponential discounting model have rarely been discussed in the literature. we explore its ability to accommodate behavioural regularities by searching both its parameter space and a reasonable stimulus space (see Table A2). This search particularly identifies its ability to accommodate the delay effect and the common difference effect (and their reversals).

Table A2 Search dimensions and domains for the double-exponential discounting model.

Dimension	Search domain	
	Delay effect	Common difference effect
β	(0, 1)	(0, 1)
δ	(0, 1)	(0, 1)
ω	(0.5, 1)	(0.5, 1)
Δt	(0, 50)	(0, 50)
t_S	(0, 10)	[0, 10]
t_{S2} - t_S		(1, 50)

Note. Parameter Space Partitioning search involves three parameters (β, δ and ω) and two (for the delay effect) or three (for the common difference effect) stimulus dimensions. Δt represents the interval between SS and LL and it is kept constant across pairs for the common difference effect. For the common difference effect, t_S and t_{S2} are the SS delay of the early- and later-onset pairs respectively

For the delay effect, we compare the implied (exponential) discount factors for the smaller delay, t_S, and the counterpart from the larger delay, t_L (denoted as $t_S + \Delta t$ in Table A2) given the double-exponential discounting model and a set of parameter thereof. Specifically, with a set of parameters (β, δ and ω), the implied discount factors

for both the interval between 0 and t_S and the interval between 0 and t_L are estimated. Particularly, the implied per-period discount factor for a delay t is given as $\dot{\delta}_{0\to t} = \left(\omega\beta^t + (1-\omega)\delta^t\right)^{1/(t-0)}$. The results suggest that the implied discount factor is constantly larger for the longer delay (t_L) than for the shorter one (t_S) given any values from the parameter space and stimulus space, producing a delay effect.

For the common difference effect and its reversal, we compare the implied discount rate over the interval Δt with an earlier onset, t_S, and that from the later onset, t_{S2}. The implied per-period discount factor over interval Δt with onset t is given as $\dot{\delta}_{t\to t+\Delta t} = \left(\frac{\omega\beta^{t+\Delta t} + (1-\omega)\delta^{tt+\Delta t}}{\omega\beta^t + (1-\omega)\delta^t}\right)^{1/\Delta t}$. The results suggest that the implied per-period discount factor over interval Δt is constantly larger for the larger onset delay (t_{S2}) than for the smaller onset delay (t_S), producing a common difference effect.

Intertemporal choice heuristics (ITCH) model. Weak stochastic transitivity (WST) does not necessarily hold for the Intertemporal choice heuristics (ITCH) model. Thus, we use Parameter Space Partitioning to identify if it can accommodate two violations of the weak stochastic transitivity: subadditivity and superadditivity.

Following Scholten et al. (2014), I used a tuple of three intertemporal choice items for this search. The three options are denoted as SS (x_S, t_S), MM (x_M, t_M) and LL (x_L, t_L) respectively with $0 < x_S < x_M < x_L$ and $0 \leq t_S < t_M < t_L$. We simulate binary choices between any of the three pairs of options based on a sample of parameters from the ITCH model. Intertemporal choice is subadditive if SS is preferred to MM, MM is preferred to LL but LL is preferred to SS (subadditivity: SS $\succ$ MM, MM $\succ$ LL, SS $\prec$ LL). Intertemporal choice is superadditive if LL is preferred to MM, MM is preferred to SS but SS is preferred to LL (superadditivity: SS $\prec$ MM, MM $\prec$ LL, SS $\succ$ LL). To facilitate the search process, the SS outcome (x_S) and the SS delay (t_S) are set to 1 respectively. The search process involves both the parameter space (κ, w_t and w_x) and the stimulus space (t_M, t_L, x_M and x_L) as shown in Table A3. From the search of both the parameter and stimulus spaces, we identify that the ITCH model could accommodate both subadditivity and superadditivity in terms of the violation of WST. Table A3 also gives examples of parameters and stimuli when the model accommodates subadditivity and superadditivity respectively. However, because this feature is unsystematic for the ITCH model, it is not featured in Table 2.3.

Table A3 Search dimensions, domains and examples for the intertemporal choice heuristics model (ITCH).

Dimension	Search domain	Examples	
		Subadditivity	Superadditivity
κ	(e^{-20}, e^{20})	6.062	0.099
w_t	(0, 1)	0.109	0.878
w_x	(0, 1)	0.608	0.064
t_M	(0, 10)	5.029	5.780
t_L	(10, 50)	28.031	44.830
x_M	(0, 10)	9.952	1.743
x_L	(10, 50)	45.924	33.511

Note. Examples of ITCH accommodating subadditivity and superadditivity respectively were provided by the search algorithm.

Reference

Pitt, M. A., Kim, W., Navarro, D. J., & Myung, J. I. (2006). Global model analysis by parameter space partitioning. *Psychological Review*, *113* (1), 57–83. doi:10.1037/0033-295X.113.1.57

Appendix 2D

Relative nonadditivity

In combination with the Luce specification, all attribute-based models (except for the proportional difference model (Cheng & González-Vallejo, 2016; PD)), can accommodate relative nonadditivity in terms of the violation of the product rule. The cases for the full tradeoff model (TM) and basic tradeoff model 1 (Scholten & Read, 2013; BTM_{SR}) have been illustrated in Scholten et al. (2014, pp.420) with a numerical example. This appendix shows this feature for the five attribute-based models with a formal proof. It also shows why the PD model (Cheng & González-Vallejo, 2016; PD) do not have this feature.

Proof.

Following Scholten et al. (2014), I used a tuple of three options to examine whether each attribute-based model could produce the pattern of relative nonadditivity in terms of the violation of the product rule. The three options are denoted as SS (x_S, t_S), MM (x_M, t_M) and LL (x_L, t_L) respectively with $0 < x_S < x_M < x_L$ and $0 \leq t_S < t_M < t_L$. The product rule of transitivity suggests that the odd of choosing LL in a choice between LL and SS should be the product of the odd of choosing LL in a choice between LL and MM and the odd of choosing MM in a choice between MM and SS: i.e., Ω(LL, SS) = Ω(LL, MM) × Ω(MM, SS). If Ω(LL, SS) ≠ Ω(LL, MM) × Ω(MM, SS), it means that the choice pattern is intransitive or nonadditive (violating the product rule). Specifically, the choice pattern is subadditive if Ω(LL, SS) > Ω(LL, MM) × Ω(MM, SS) and superadditive if Ω(LL, SS) < Ω(LL, MM) × Ω(MM, SS). Relative nonadditivity in the violation of the product rule means that the ratio $\frac{\Omega(\mathrm{LL,SS})}{\Omega(\mathrm{LL,MM}) \times \Omega(\mathrm{MM,SS})}$ decreases as the amounts (x_S, x_M and x_L) increase in proportion.

Combined with the Luce specification, the ratio can be written as a function of advantages functions (outcome advantage $V(.)$ and time advantage $Q(.)$):

$$\frac{\Omega(\mathrm{LL,SS})}{\Omega(\mathrm{LL,MM}) \times \Omega(\mathrm{MM,SS})} = \frac{\left(\frac{V(x_S,x_L)}{Q(t_S,t_L)}\right)^{1/\varepsilon}}{\left(\frac{V(x_M,x_L)}{Q(t_M,t_L)}\right)^{1/\varepsilon} \times \left(\frac{V(x_S,x_M)}{Q(t_S,t_M)}\right)^{1/\varepsilon}} = \left(\frac{V(x_S,x_L)}{V(x_M,x_L)\cdot V(x_S,x_M)}\right)^{1/\varepsilon} \cdot \left(\frac{Q(t_M,t_L)\cdot Q(t_S,t_M)}{Q(t_S,t_L)}\right)^{1/\varepsilon}.$$

Let $x_M = a \times x_S$ and $x_L = b \times x_S$ $(1 < a < b)$ and $g(x_S) = \frac{V(x_S,x_L)}{V(x_M,x_L)\cdot V(x_S,x_M)} = \frac{V(x_S,bx_S)}{V(ax_S,bx_S)\cdot V(x_S,ax_S)}$. It is obvious that $\frac{\Omega(\mathrm{LL,SS})}{\Omega(\mathrm{LL,MM}) \times \Omega(\mathrm{MM,SS})}$ is a monotonically increasing

function of $g(x_S)$. Thus, for each model, to determine whether it accommodate relative nonadditivity in terms of the violation of the product rule, we need only to know if $g(x_S)$ is a monotonically decreasing function of x_S. If, for a model, $g(x_S)$ is a monotonically decreasing function of x_S, then $\frac{\Omega(\mathrm{LL,SS})}{\Omega(\mathrm{LL,MM}) \times \Omega(\mathrm{MM,SS})}$ decreases as x_S increases, showing that the model, in combination with the Luce specification, produces a pattern of relative nonadditivity in the violation of the product rule. Otherwise, this model does not reliably predict a pattern of relative nonadditivity in terms of the violation of the product rule.

Full Tradeoff Model (Scholten et al., 2014; TM). For TM, $g(x_S) = \frac{\gamma(\log(1+\gamma b x_S)-\log(1+\gamma x_S))}{(\log(1+\gamma b x_S)-\log(1+\gamma a x_S))\cdot(\log(1+\gamma a x_S)-\log(1+\gamma x_S))}$

The derivative $g'(x_S) =$

$$-\frac{\gamma^2\left(\frac{b-a}{(b\gamma x_S+1)(\log(1+a\gamma x_S)-\log(1+b\gamma b x_S))^2}+\frac{a-1}{(\gamma x_S+1)(\log(1+\gamma x_S)-\log(1+a\gamma x_S))^2}\right)}{a\gamma x_S+1} < 0 \text{ for any}$$

parameter and stimulus values in the domain. Thus, $g(x_S)$ is a monotonically decreasing function of x_S and the full tradeoff model can accommodate relative nonadditivity in terms of the violation of the product rule.

Basic Tradeoff Model 1 (Scholten & Read, 2013; BTM$_{SR}$). $g(x_S)$ for BTM$_{SR}$ is the same as $g(x_S)$ for the full tradeoff model above. Thus BTM$_{SR}$, in combination with the Luce specification, can also accommodate relative nonadditivity in terms of the violation of the product rule.

Basic Tradeoff Model 2 (Dai & Busemeyer, 2014; BTM$_{DB}$). $g(x_S)$ for BTM$_{DB}$ is:

$$g(x_S) = \frac{(bx_S)^\gamma - x_S{}^\gamma}{((bx_S)^\gamma-(ax_S)^\gamma)\cdot((ax_S)^\gamma-x_S{}^\gamma)}.$$

The derivative $g'(x_S) = -\frac{\gamma\left(\frac{1}{(bx_S)^\gamma-(ax_S)^\gamma}+\frac{1}{(ax_S)^\gamma-x_S{}^\gamma}\right)}{x_S} < 0$ for any parameter and stimulus values in the domain. Thus, BTM$_{DB}$ can accommodate relative nonadditivity in terms of the violation of the product rule.

Intertemporal choice heuristics model (Ericson et al., 2015; ITCH). $g(x_S)$ for the ITCH model is:

$$g(x_S) = \frac{(1-w_x)(b-1)x_S+2w_x\frac{b-1}{b+1}}{\left((1-w_x)(b-a)x_S+2w_x\frac{b-a}{b+a}\right)\left((1-w_x)(a-1)x_S+2w_x\frac{a-1}{a+1}\right)}.$$

$g'(x_S)=$

$$-\frac{(a+1)(b-1)(1-w_x)(a+b)\left((a+1)(b-1)(1-w_x)^2(a+b)x_S{}^2+4(a+1)(1-w_x)(a+b)w_x x_S+8aw_x{}^2\right)}{(a-1)(b+1)(b-a)\left((a+1)(a+b)(1-w_x)^2x_S{}^2+2(2a+b+1)(1-w_x)w_x x_S+4w_x{}^2\right)^2} < 0$$

for any parameter and stimulus values in the domain. Thus, the ITCH model can accommodate relativity nonadditivity in terms of the violation of the product rule.

Difference-ratio-interest-finance-time model (Read et al., 2013; DRIFT). For the DRIFT model,

$g(x_S) =$

$$\frac{\left((1-w_I)(1-w_R)(b-1)x_S+w_R(b-1)+w_I\left(b^{\frac{1}{t_L-t_S}}-1\right)\right)}{\left((1-w_I)(1-w_R)(b-a)x_S+w_R\frac{(b-a)}{a}+w_I\left(\left(\frac{b}{a}\right)^{\frac{1}{t_L-t_M}}-1\right)\right)\left((1-w_I)(1-w_R)(a-1)x_S+w_R(a-1)+w_I\left(a^{\frac{1}{t_L-t_S}}-1\right)\right)}.$$

The derivative, $g'(x_S)$, becomes intractable and we do not have an analytical solution to determining whether $g'(x_S)$ is constantly negative given any parameter and stimulus values. Instead, we used the *fmincon* function in Matlab (with the "interior-point" optimisation algorithm) to determine whether $g'(x_S)$ is constantly negative given any parameter and stimulus values. Extensive search of the parameter and the stimulus spaces suggests that $g'(x_S)$ for the DRIFT model is constantly negative given any parameter and stimulus values. Thus, the DRIFT model can also accommodate relative nonadditivity in terms of the violation of the product rule.

Proportional difference model (PD). For the proportional difference model, $g(x_S) = \frac{\left(\frac{b-1}{b}\right)}{\left(\frac{b-a}{b}\right)\cdot\left(\frac{a-1}{a}\right)}$. It is obvious that x_S is cancelled out in $g(x_S)$. So this model is silent on relative nonadditivity.

References

Cheng, J. & González-Vallejo, C. (2016). Attribute-wise vs. alternative-wise mechanism in intertemporal choice: Testing the proportional difference, trade-off, and hyperbolic models. *Decision, 3*(3), 190-215. doi:10.1037/dec0000046

Dai, J. & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, *143*(4), 1489–1514. doi:10.1037/a0035976

Ericson, K. M. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier

or later? Simple heuristics explain intertemporal choices better than delay discounting does. *Psychological Science*, *26*(2), 826–833. doi:10.1177/0956797615572232

Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An analysis of outcome framing in intertemporal choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 573–588. doi:10.1037/a0029177.

Scholten, M. & Read, D. (2013). Time and outcome framing in intertemporal tradeoffs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39* (4), 1192–1212. doi:10.1037/a0031171

Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes by time or against time? Evaluation rules in intertemporal choice. *Cognitive Science*, *38*(3), 399–438. doi:10.1111/cogs.12104

Appendix 3A

Cognitive Reflection Test

Please answer the following questions by checking the answer you think is correct.

1. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?
$0.05
$0.10
$0.50
$1.00

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
5 minutes
100 minutes
20 minutes
500 minutes

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
47 days
24 days
36 days
29 days

Note. The order of the options in each CRT question was shuffled.

Appendix 3B

Demographic information questions

Please provide some demographic information below. All information you provide will be kept anonymous and confidential and will only be used for group analysis

1.What is your gender?

Female

Male

2.What is your age in years? _____

3.What is the highest education degree you have earned

Less than high school

High school

Bachelor's degree or equivalent

Master's degree or equivalent

Doctoral degree or equivalent

Others

4.Have you taken any university-level courses in the following subjects? (please check all that apply)

Economics

Psychology

Finance

Mathematics

Statistics

None of the above

5.What is your race?

White/Caucasian

African American

Hispanic

Asian

Native American

Pacific Islander

Others

6.What is your current employment status?

Full-time employed

Part-time employed

Self-employed

Unemployed

Retired

Full-time student

Others

7. Please provide a rough estimate of the total combined income of all members of your household in the last year.

Less than $20,000

$20,000 - $39,999

$40,000 - $59,999

$60,000 - $79,999

$80,000 or more

Prefer not to answer

Appendix 3C

Results relating to cognitive reflection

Aside from attention manipulation, the two experiments also measured participants' level of cognitive reflection with the Cognitive Reflection Test (CRT; Frederick, 2005). To test the potential effects of cognitive reflection and its interaction effect with manipulated attention on intertemporal preference, participants in both experiments were further divided into two sub-groups according to their performance in the Cognitive Reflection Test (CRT). Those who correctly answered all three questions in CRT are considered high in cognitive reflection and others low in cognitive reflection.

Table A4 Parameter estimation from models involving cognitive reflection.

Fixed-effect	Parameter	Experiment 1	Experiment 2
Attention	β_1	**0.532 [0.452, 0.613]**	**-0.177 [-0.256, -0.100]**
Background Contrast	β_2	**0.062 [0.052, 0.072]**	**-0.057 [-0.067, -0.048]**
Attention × Background Contrast	β_3	-0.009 [-0.033, 0.015]	**-0.028 [-0.055, -0.003]**
CRT	β_4	**0.743 [0.664, 0.824]**	**0.641 [0.563, 0.720]**
Attention × CRT	β_5	**-0.652 [-0.811, -0.493]**	**-0.292 [-0.451, -0.136]**
Background Contrast × CRT	β_6	**0.034 [0.017, 0.050]**	-0.005 [-0.021, 0.012]
Attention × Background Contrast × CRT	β_7	-0.015 [-0.048, 0.017]	0.014 [-0.047, 0.019]

Note. Attention is an effect-coded variable: -0.5 = delay-focus; 0.5 = outcome-focus. *Background contrast* ranges from -1 (dominated by all other LL options) to 1 (dominating all other LL options). *CRT* is also an effect-coded variable: -0.5 = low cognitive reflection; 0.5 = high cognitive reflection. The estimates outside the brackets are the median (Md) and the estimates in the brackets are the 95% High Density Intervals (95% HDIs) of the 150,000 samples from the posterior distribution. The 95% HDIs of the cells in boldface do not cross 0.

Table A4 presents the results from the mixed-effect Bayesian models involving cognitive reflection from the two experiments respectively. First of all, comparisons with Table 3.1 (pp. 69 in the thesis) and Table 3.2 (pp. 73 in the thesis) suggest that involving cognitive reflection in the model does not change the results regarding the attention effects and the background contrast effects. Thus this is a further robustness check of the main findings from Chapter 3. Second, consistent with findings from Frederick (2005), the main effect of cognitive reflection suggests that participants with high cognitive reflection were less impatient than those with low cognitive reflection in both experiments.

Table A5 Average proportions of LL choices in the two experiments, segregated by cognitive reflection groups.

EXPERIMENT 1: ATTENTION TO LL COMPONENTS		
Cognitive reflection	Attention condition	
	Delay-focus	Outcome-focus
High	39.7% (*N* = 42)	44.1% (*N* = 53)
Low	21.4% (*N* = 48)	35.7% (*N* = 43)
Difference between LL proportions of high- and low-cognitive-reflection groups	**18.3%**	**8.4%**
EXPERIMENT 2: ATTENTION TO SS COMPONENTS		
Cognitive reflection	Attention condition	
	Delay-focus	Outcome-focus
High	50.4% (*N* = 43)	45.5% (*N* = 47)
Low	37.1% (*N* = 53)	37.2% (*N* = 50)
Difference between LL proportions of high- and low-cognitive-reflection groups	**13.3%**	**8.3%**

Third, the effect of cognitive reflection on intertemporal choice differs across attention conditions. In particular, in Experiment 1 outcome-focus, compared with delay-focus, attenuates the effect of cognitive reflection on impatience in intertemporal preference. The top panel of Table A5 presents an intuitive illustration of this effect. In the delay-focus condition, the difference between the LL proportion of low-cognitive-reflection participants and the LL proportion of high-cognitive-

reflection participants is 18.3%, whereas the difference in the outcome-focus condition is only 8.4%. Similarly, in Experiment 2, outcome-focus also reduces the effect of cognitive reflection on impatience in intertemporal preference. As shown in the bottom panel of Table A5, the difference between the overall LL proportion from the participants with high cognitive-reflection and the overall LL proportion from the participants with low cognitive-reflection in the outcome-focus condition (8.3%) is smaller than the difference in the delay-focus condition (13.3%). Collectively, the moderation of manipulated attention on the effect of cognitive reflection on intertemporal preference suggested that excessive impatience among people with low cognitive reflection could be, at least, partially attributable to inattention or a lack of attention to outcomes and more attention to delays.

References

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19* (4), 25–42. doi:10.1257/089533005775196732

Appendix 4A

Consideration of Future Consequences scale

We will ask you to evaluate how much a number of statements are characteristic of you. Please click on the appropriate box for each statement.

1. I consider how things might be in the future, and try to influence them with my behaviour.

Extremely uncharacteristic
Somewhat uncharacteristic
Uncertain
Somewhat characteristic
Extremely characteristic

2. I often do things to achieve outcomes that may not result for many years.

Extremely uncharacteristic
Somewhat uncharacteristic
Uncertain
Somewhat characteristic
Extremely characteristic

3. I only act to satisfy immediate concerns. The future will take care of itself.

Extremely uncharacteristic
Somewhat uncharacteristic
Uncertain
Somewhat characteristic
Extremely characteristic

4. My behaviour is only influenced by the effects it will have in a matter of days or weeks.

Extremely uncharacteristic
Somewhat uncharacteristic
Uncertain
Somewhat characteristic
Extremely characteristic

5. Convenience is a big factor in my choices and actions.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

6. I will sacrifice immediate happiness to achieve future outcomes.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

7. It is important to pay attention to warnings even when they are about outcomes that will not occur for many years.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

8. It is better to do things with large delayed effects than things with small immediate effects.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

9. I usually ignore warnings about possible future problems. They can be resolved later, before they reach crisis level.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

10. Sacrifices here and now are usually unnecessary. The future can be dealt with at a later time.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

11. I only act to satisfy immediate concerns. I will take care of future problems at a later date.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

12. My day to day work has immediate affects, so it is more important to me than behaviour that has distant affects.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

13. I always think about how what I decide might affect me in the future.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

14. My behaviour is usually affected by future consequences.

Extremely uncharacteristic

Somewhat uncharacteristic

Uncertain

Somewhat characteristic

Extremely characteristic

Appendix 4B

Screening items for Experiment 1 in Chapter 4

Condition	Option A	Option B
No-FED	£37 today	£25 in 17 days
No-FED	£84 today	£70 in 95 days
No-FED	£62 today	£60 in 169 days
FED	£37 in 100 days	£25 in 117 days
FED	£84 in 100 days	£70 in 195 days
FED	£62 in 100 days	£60 in 269 days

Appendix 4C

Demographic information questions

Please provide some demographic information below. All information you provide will be kept anonymous and confidential and will only be used for group analysis

1. What is your gender?
Male
Female
Others/Prefer not to say (3)

2. What year were you born? _____

3. Which country do you come from? _____

4. What is the highest level of education you have completed?
Less than secondary school
Secondary school
Some university
Bachelor's degree
Master's degree
Doctoral degree
Professional degree (JD, MD)

5. What is your employment status?
Full-time
Part-time
Unemployed
Other

6. Please indicate your current family structure.
Single without children
Single with children

Married without children

Married with children

Life partner without children

Life partner with children