PALACKÝ UNIVERSITY OLOMOUC
FACULTY OF ARTS

The Price of Now:
Economic Implications of Propensity for Immediate Gratification

Bachelor Thesis

Author: Marek Hudeček
Supervisor: Mgr et. Mgr Jan Stoklasa, Ph.D
Olomouc 2024

Declaration

I hereby solemnly swear that I have independently prepared the bachelor's thesis on the topic *The Price of Now: Economic Implications of Propensity for Immediate Gratification* under the professional supervision of my thesis supervisor Mgr. et Mgr. Jan Stoklasa, Ph.D, and have included in it all the used sources and literature.

In Přerov, date: 18.4.2024 Signature: 

Acknowledgement

I wish to express my gratitude to Mgr. et Mgr. Jan Stoklasa, Ph.D. Without his support, this thesis would most likely not exist.

myopia /maɪˈəʊpiə/ *noun*. **1.** The condition of being unable to see things clearly when they are far away, synonyms: short sight, short-sightedness

2. The state of being unable to see what the results of a particular action or decision will be; the failure to think about anything outside your own situation

Oxford Advanced Learner's Dictionary (Hornby et al., 2020)

Contents

Abstract

Myopic time-inconsistent preferences leading to self-control failure, or preference reversals in intertemporal choice where the larger, later reward is preferred over smaller, sooner reward at a distance but neglected when the smaller reward is immediately available have long been the subject of interest by philosophers, theologists, and scientists from many fields, such as psychology, economics, or neurophysiology. This thesis provides an introductory overview of contemporary normative and descriptive economic theories surrounding intertemporal choice. It presents empirical findings from both human and animal research to elucidate differences in temporal discounting across individuals, goods and time horizons to identify some of the factors which influence steepness of discounting and preference reversals. It also investigates the topic from the perspective of psychology and neuroscience, which provide alternative explanations of self-defeating behavior which do not build on an internal discount function, with special consideration given to multiple-process models. Using a framework by Story et al. 2014, I investigate a possible mechanism of interaction between habitual and goal-directed learning systems through the perspective of reinforcement learning, which could explain the habituation and subsequent overconsumption of hedonic goods, i.e., goods which provide immediate pleasure and have delayed costs. The model-free (habitual) system explains our seemingly inexplicable inability to forego immediately rewarding goods despite their long-term costs, while the model-based (goal-directed) system explains our ability to change discount rates by focusing on different information. I then review some of the proposed strategies for long-term utility maximization, as well as discuss possible directions of future intertemporal choice research.

Keywords: intertemporal choice, temporal discounting, delay discounting, hyperbolic discounting, behavioral economics, utility maximization, self-control, addiction

Introduction

Every day, we make decisions. What to eat, what to wear, what to do, what to spend our money on. More often than not, making a decision to pursue one goal involves the sacrifice of not being able to pursue another goal. Our desires, ultimately, are unlimited[1], yet our time, money, and energy are not. This is the fundamental essence of economics, the study of how people use their limited resources to best satisfy their needs and desires **(Heshmat, 2015a, p. 218)**. Some goals are even direct opposites of each other, such as the desire to drink in a person who also wishes to be sober. **George Ainslie (1992)** coined the term picoeconomics (or micromicroeconomics) to describe the property of a utility-maximizing brain in which multiple interests strive for attention, generating an internal marketplace of sorts, where the interest which can offer the highest reward outbids the others.

Comparing the relative desirability of two options is simple enough, if they are immediately available. If we are to choose only on the dimension of amount, we should prefer a larger over a smaller amount, assuming we are rational in the sense that it is rational to maximize utility, and that utility increases with amount (irrespective of the exact shape of the utility function).

Let's add another dimension to the equation – time. As will be discussed later in the thesis but the reader probably already knows, people (and animals) have positive time preference towards rewards **(Kalenscher and Pennartz, 2008; Story et al., 2014; Callender, 2021)**. What this means is, that all other things staying the same, we prefer to receive a reward of the same amount sooner rather than later. In other words, we discount delayed rewards, and the extent to which a reward is discounted is increased as the delay increases. Exponentially discounted utility model proposed by **Samuelson (1937)** assumes that delayed utility ought to be discounted exponentially, i.e., by the same percentage for every constant time step. This model was largely adopted by neoclassical economists as a normative description of rational behavior. Furthermore, provided that people represent instantaneous utility derived from choices on a single dimension, another assumption often made by economists, we have no reason other than to assume

[1] Or, at the very least, orders of magnitude higher than what our available resources allow us to fulfill.

that utilities from different goods have the same discount rate. What this means is that traditional economists often assume that intertemporal (involving different time periods) preferences are consistent in time, or time-consistent **(O'Donoghue & Rabin, 2000)**. Yet, an immense amount of evidence has been amassed which suggests the exact opposite **(e.g. Vuchinich & Heather, 2003; Kalenscher & Pennartz, 2008; Urminsky & Zauberman, 2016)**. Both people and animals have repeatedly shown the tendency to switch preferences towards immediately available rewards after preferring the larger, more delayed reward from a distance. This means that when plotting how discounted values change in time, i.e., the assumed discount curves, at some point, the curves cross **(Ainslie, 2001, p. 31)**.

The real life implication of this is the phenomenon of self-defeating behavior, which has puzzled people for millennia **(Ainslie, 2001, pp. 7–11; O'Donoghue & Rabin, 2000)**. We call this behavior self-defeating, because as smaller, closer rewards become readily available, we shift our preference towards them and in doing so, we forego the chance to obtain a higher reward, which we later regret. Some examples of such behavior include eating dessert while trying to lose weight, smoking while promising ourselves that we would stop, procrastinating, failing to save or invest money, or, in extreme cases, crippling addiction to drugs such as heroin or alcohol. Consequences of such behavior can have devastating effects on both nations and individuals.

The goals of this thesis are to firstly provide an introductory overview of the contemporary research surrounding myopic self-defeating behavior from several perspectives. Economic theories of intertemporal choice will be assessed with regards to the consistency of their predictions with real-world evidence. Individual differences in temporal discounting between different people and between different goods will be investigated. Going beyond the field of traditional economics, I will also include an account of behavioral and cognitive patterns associated with self-defeating behavior along with possible mechanisms of their origin. In particular, I will evaluate time-inconsistent preferences through the perspective of multiple-process models, incorporating research from the fields of psychology, neuroscience and reinforcement learning.

The second goal is to identify factors which influence intertemporal choice evaluation, as well as to provide strategies which are designed to combat the

overvaluation of immediate rewards and neglect of distant rewards, and which therefore improve long-term utility maximization of individuals. Furthermore, I will discuss possible directions of intertemporal choice research both in the realm of economics and behavioral economics, as well as psychology and public choice.

The value of the thesis will depend on its reader. As a whole, it helps understand self-defeating behavior, as well as provide strategies for those who wish to mitigate it. An economist could find value in the additional perspective of psychology, neuroscience and reinforcement learning. Addiction specialists can benefit from viewing the phenomenon of addiction and other self-defeating behaviors through the perspective of behavioral economics, as well as by understanding the various factors which influence deliberate choice. Some theories of addiction largely ignore or even undermine the role of conscious, deliberate choice in addiction. While the thesis shows that certain physiological, environmental and psychological factors do increase the risk of addiction, and that in certain conditions, recovery may be impossible due to insufficient cognitive resources, deliberate choice is an undeniable factor in addiction and recovery, and treating those afflicted as powerless could hurt them in the long term. Last but not least, it can be of tremendous value to lawmakers and public policy makers, as it can help them understand how choice architecture, laws, taxes and public resource allocation influences the wellbeing of individuals, which can help them utilize utility-maximization strategies on an organizational level.

My personal motivation to write this thesis is to gain insight into the topic of self-defeating behavior and self-control failure, because I myself am no stranger to procrastination or impatience. I believe that by learning about the topic, it will at some point help me or somebody else to change the course of their lives for the better. I also have no reason to hide the fact that I find the topic fascinating, and see this thesis, if nothing more, as a stepping stone for my future personal, academic and professional development. I sincerely hope that reading this thesis will be of value to you.

Utility, choice and rationality

How do we even make choices? After all, when having to sacrifice one option for the other, we often have to decide between choices which have seemingly nothing in common. How can we prefer to spend $10 now for a pizza over spending those same $10 for a movie ticket tomorrow, or over saving them towards purchasing a bike? Most likely, we will choose by a feeling, a semi-conscious calculation in which we evaluate which of the choices available to us will make us the most pleased. In other words, we try to condense the seemingly infinite array of variables that are relevant to the choice, such as how hungry we are, whether the movie will be available next week, whether we already have an older bike, and so on, into one dimension, which would allow us to directly compare the choices. In the introduction, I used the term utility. This utility is the term which economists use to denote this measure of desirability. To be exact, neoclassical economists have used the argument that it is *as if* the brain computed some kind of measurement of desirability called utility. Another term I used, rationality, isn't universally defined across sciences, however, in this thesis, I consider rationality a deliberate effort to maximize utility, i.e., the wellbeing of an individual. Neoclassical economists, working under this assumption that all human decisions are serving the rational effort to maximize utility, have developed a mathematical foundation known as the utility theory, which provides a prescriptive (normative) framework of rational decision making **(Glimcher et al., 2005)**.

However, what followed was a series of empirical studies and experiments, which showed that people systematically depart from the predictions of utility theory on rational behavior **(Kahneman & Tversky, 1979)**. This suggests that either humans aren't fully rational utility maximizers, or that the assumptions underlying neoclassical utility theory were wrong or incomplete **(Glimcher et al., 2005)**. Different approaches were employed to deal with the issue of people not acting according to the predictions of utility theory for rational behavior. Herbert A. Simon coined the term bounded rationality, which describes the limitations of the brain's computational capacity and the ability to evaluate all alternatives, and observed that people's ability to maximize utility according to standard utility theory is bounded to only some cases **(Simon, 1990)**. Others argued that

standard theory requires modifications or different approaches **(Kahneman & Tversky, 1979).**

An interesting change in perspective came from the field of neuroscience. Some authors have argued that the brain must actually have some kind of common weight by which qualitatively different choices are evaluated for an orderly choice to be possible **(Shizgal & Conover, 1996).** Therefore, there emerged a field of neuroeconomics, which aims to make both prescriptive and normative economic theories consistent with actual economic computations done by the human brain **(Glimcher et al., 2005).** Advances in neuroeconomics showed that the brain actually does make a computation of desirability for each available course of action **(Glimcher et al., 2005; Shizgal & Conover, 1996).** Neuroeconomics also redefined the question of rationality.

> First, under many conditions, conditions under which choice appears rational, this desireability encoded by the neurons of the brain very closely approximates expected utility. Second, under conditions in which choice behavior is poorly predicted by rational choice models, these neural representations still encode the desireability of each course of action, although under these conditions desirability and expected utility are of necessity not identical. **The available data suggest that the neural decision-making process is always rational with regard to these internal representations of desireability.** When choosers deviate from rationality it is this physiological encoding of desirability, which we refer to as *physiological expected utility*, that departs from neoclassical theory **(Glimcher et al., 2005, p. 7)**

In other words, our brains seem to be rational in the sense that they are capable of choosing a course of action which has the highest *physiological* expected utility, but under some circumstances, this physiological expected utility is different from expected utility predicted by the expected utility theory. Experiments on rats by **Shizgal and Conover (1996)** also suggest that ultimately, desirability of a reward seems to be one-dimensional, determined by the level of electrical stimulation of reward-related neurons multiplied by the number of those neurons, and that this desirability is relative to desirabilities of alternative courses of action available.

We must also discern between different kinds of utility, as there is more than just one kind of utility **(Heshmat, 2015a, p. 32)**.

decision utility, also **predicted** or **expected utility** denotes the utility that a person expects to gain from a choice. It is the weight that gets put on the scale before a decision is made to take an action. In an animal experiment, it would be the amount of food that the animal expects to receive after pressing a lever, or more accurately, the utility that the animal expects to derive from the food that it expects to receive.

experienced utility is the "hedonic impact of the reward that is actually experienced when it is finally gained **(Heshmat, 2015a, p. 32)**". Simply put, it is the actual reward received after a decision. For the animal, it would be the actual utility (electrical stimulation of reward neurons) derived from the amount of food that it receives in a single instance of pressing the food lever.

remembered utility is the utility associated with the choice after it was made. If we assume that it stays constant, we can say that it is the updated expected reward of an action after it is made. Because we live in a world where experienced utility frequently varies from expected utility, because our information is never complete – such as ordering a meal in a restaurant that we expect to taste a certain way, and finding that it tastes better or worse than we expected – with subsequent orders of the same meal, our expected and remembered utilities become an aggregate of a stream of experiences **(Heshmat, 2015a, p. 33)**.

However, this should not imply that the process of physiological calculation of desirability of an action is simple, centralized, or even completely understood. While it is possible that in a given moment, available choices are ultimately represented on a single dimension of desirability, I must highlight that even if that is the case, we are still looking at the final part of the process of choice evaluation. To fully capture economic decision making using neurobiology and to explain observed phenomena through neurobiological processes, we must not only uncover the final mechanism of simply choosing the most desirable choice at the moment, but also the factors which influence the desirability of an outcome, the way in which they are processed and by which part of the system, how those systems cooperate, what influences them, and how this process changes over time. For example, there are several areas of the brain which process information about reward

amount **(Kalenscher & Pennartz, 2008).** Also, several regions of the brain were linked to the representation and perception of time, and numerous physiological mechanisms were proposed which would explain how passage of time is measured or approximated **(Kalenscher & Pennartz, 2008).** Desirability of a choice can also change in time based on the interaction of different forms of learning systems **(Story et al., 2014)**, and even the perception of time can be changed by numerous factors **(Sayette et al. 2005; Li et al., 2022)**.

This chapter shows that decision making, including intertemporal choice, can never be fully understood without understanding the underlying biological mechanisms which govern it. Only an interdisciplinary approach has the chance of us seeing the full picture.

Intertemporal choice

We are now getting to the main topic of the thesis – how does delaying the experienced reward influence the immediate value of the delayed reward? In other words, how is the present value of future expected utility discounted as a function of the delay length?

People and animals have repeatedly shown to have positive time preference **(e.g. Kalenscher & Pennartz, 2008; Story et al., 2014; Callender, 2021)**. This means that a reward is worth more when it is delivered sooner compared to the same reward delivered later. This makes biological sense considering the risks and uncertainty associated with the environment in which we evolved, as well as the fact that we are mortal, and that some appetites, such as for food or sleep, have a limited time to be satisfied. From a strictly financial sense, it also seems rational to discount money at the very least at the same rate that we could obtain from a risk-free investment – why should receiving 100 dollars in a year be worth paying 100 dollars to us now, if we could receive 105 dollars after the same delay by putting 100 dollars into a savings account?

If people have positive time preference and therefore discount delayed rewards, for a delayed reward to be preferred over immediately available reward, its discounted expected utility must be larger than the expected utility of the immediate reward.

Important note is that in some cases, people seem to prefer the same rewards later rather than sooner. This, at first sight, seems like a violation of positive time preference. The answer is a bit more nuanced. People seem to derive pleasure (and therefore utility) from merely anticipating a reward **(Ainslie & Monterosso, 2003)**. Delaying the gain of a reward can also increase its utility at the point of obtainment by increasing appetite for the reward. A pizza will have different desirability and therefore expected utility or reward value based on many factors such as whether we are hungry or not. The pleasure of anticipation can be commonly observed in romantic relationships. Just as people derive positive utility from anticipation of some rewards, they can derive negative utility from certain negative outcomes, such as monetary losses, pain, or punishments – this negative utility is commonly referred to as dread. This causes the preference for negative outcomes of the same size to occur sooner rather than later. Such nuances are rather difficult to account for, the literature that I reviewed therefore, while acknowledging this

phenomenon, focuses on intertemporal preferences for rewards and costs bound to a single point in time.

Exponential discounted utility theory

Although the topic of time preference and intertemporal choice was studied even by ancient philosophers, the need to quantify it and to set a normative standard came after the emergence of complex financial instruments **(Callender, 2021)**.

After expected utility theory was developed, **Paul Samuelson (1937)** added time preferences to his framework, in which he assumed that the discount rate of future utility is constant. This means that the immediate value of future expected utility is calculated:

$$V(A,t) = A \times \delta^{t},$$

where V is the immediate value, A is the expected utility at the time of consumption, δ is the discount factor representing the fraction that a delayed reward retains in its value after one additional unit of delay, and t is the delay after which the utility is expected to be received.

Figure 1.1 showcases the dependency of immediate value of a reward worth 100 units (such as utiles) at the time of consumption on delay. In this example, utility is discounted by 15% (δ=0.85) after each unit of delay, i.e., exponentially discounted. Using the Samuelson's formula, we see that if the delay is 1 unit, the immediate value of the future utility is $100 \times 0.85^{1} = 85$ utiles. With 2 units of delay, the immediate value is $100 \times 0.85^{2} = 72.25$ utiles, with 3 units it is $100 \times 0.85^{3} = 61.4125$ utiles and so on.

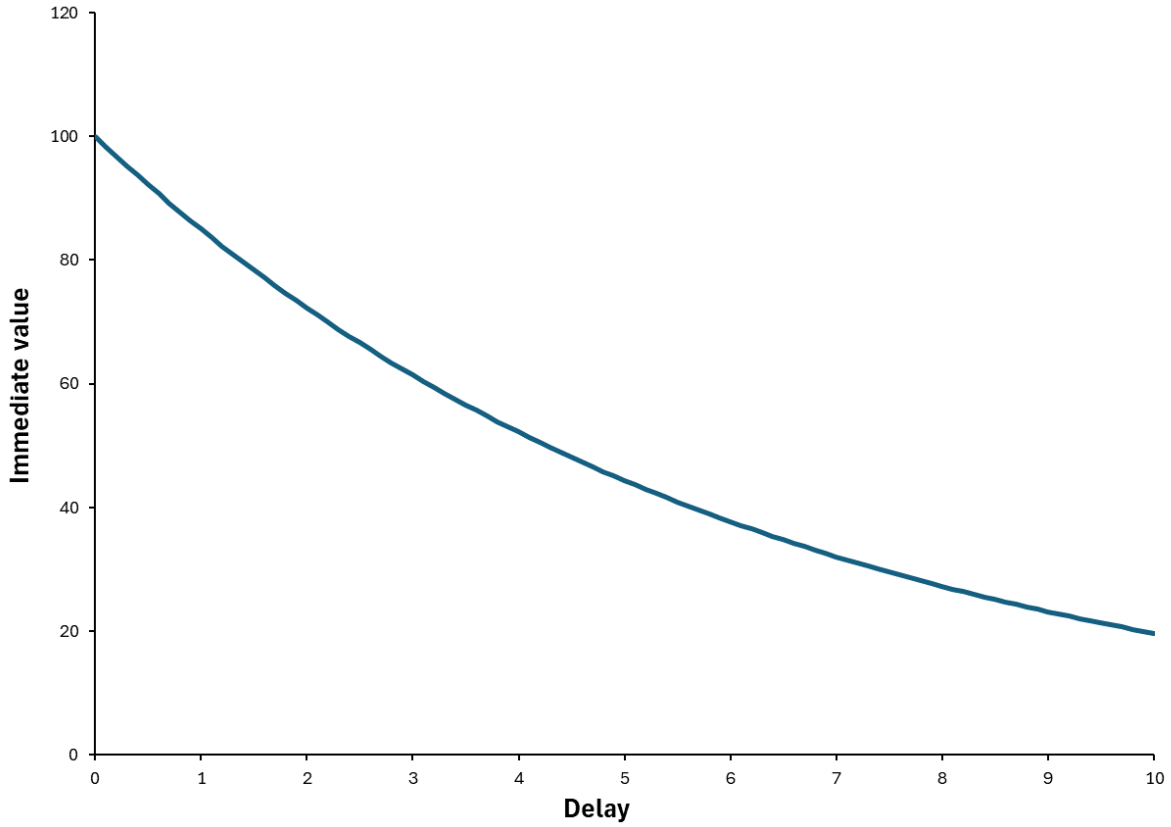

Figure 1.1: Discounted value function of a reward worth 100 units at the moment of consumption based on the length of delay. Here, the value drops 15% per each unit of delay.

Even **Samuelson (1937)** himself noted that this assumption doesn't reflect reality, despite this, exponential discounting was largely adopted to describe rational decision making. The reason behind this is that exponential discounting is the only type of discounting that is time-consistent, which means that for each interval of the same length, the same proportion of reward is discounted. Also, it is such that for any choice with delayed reward, the ratio of its discounted utility and the discounted utility of another available choice stays the same at all times that both choices are available.

Figure 1.2 effectively shows this[2]. At each point in time, the smaller future reward (50 utiles received at delay 0) is worth exactly half of the bigger future reward (100 utiles

[2] Note that unlike in Figure 1.1, it is shown that instead of decreasing reward with increasing delay, the same situation is rephrased so that the value of future utility increases as its receival moves closer in time. This will be important to illustrate the relationship between differently delayed future utilities.

received at delay 10). Again using the Samuelson's formula, we see that if the delay is 1 unit, the immediate value of the larger reward is $100 \times 0.85^1 = 85$ utiles, while the immediate value of the smaller reward is $50 \times 0.85^1 = 42.5$. With 2 units of delay, the immediate values are $100 \times 0.85^2 = 72.25$ utiles and $50 \times 0.85^2 = 36.125$ utiles respectively, and so on.

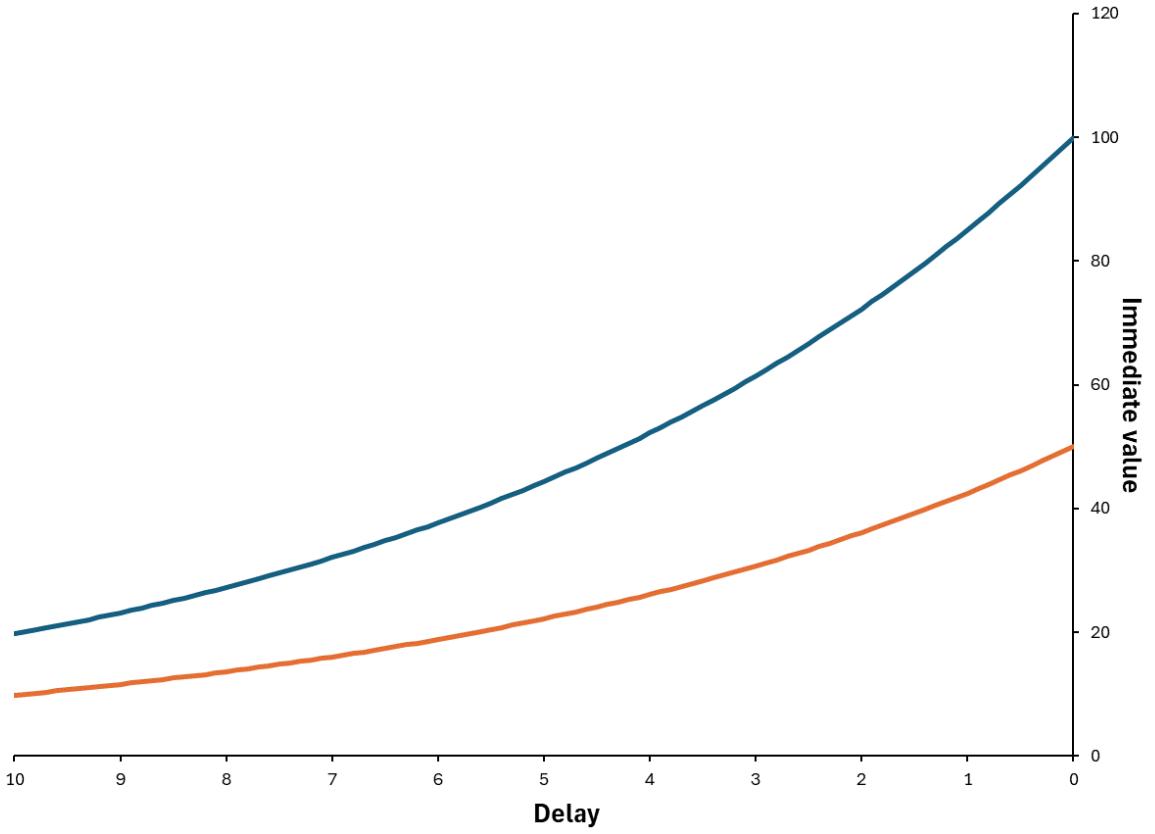

Figure 1.2: Discounted value function of rewards worth 100 and 50 units respectively at the moment of consumption based on the length of delay. The discount rate is 15% per each unit of delay.

To see why this type of discounting is deemed the only possible rational discounting in environments with known risks and no uncertainty, we must investigate the implications of non-exponential discounting. By definition, discounting so that at each point in time, we discount future utilities at the same rate, is exponential discounting. Any other discounting, therefore, will be non-exponential, and will have both moments in which we discount utility more steeply and moments where we discount less steeply than if we were exponential discounters. As will be shown later, such

variability in evaluating time predicts the occurrence of preference reversals, which leave the non-exponential discounter vulnerable to exploitation by an exponential discounter. This makes exponential discounting seem rational not only because it is consistent, but also because it is adaptive, in the sense that it is optimal in an environment with other agents who we can trade with **(Ainslie, 2001, p. 31)**.

Utility also poses one additional challenge. The law of diminishing marginal utility posits that while reward utility rises with reward amount, reward utility as a function of reward amount is not linear, instead, the function is concave, with each additional reward amount providing less utility than the previous one. Despite this, in intertemporal choice research, utility as a function of (discounted) reward value is often assumed to be linear, either implicitly or explicitly **(Andersen et al., 2008)**. **Andersen et al. (2008)** try to account for this by eliciting risk preferences together with time preferences and find significantly lower discount rates than when assuming linear utility function, however, in their paper, they assume that risk aversion mirrors the concavity of the utility function, which is correct under expected utility theory, but contradicts real-world findings about people's aversion for losses **(Kahneman & Tversky, 1979)**. **Ubfal (2016)** also reports that the majority of the studies they analyzed assumed linear utility function. Therefore, this thesis will mostly focus on discount rates of reward amounts.

Time consistency

The received view is that an agent's choices should be consistent in time under unchanged information to be deemed rational, otherwise, the agent becomes vulnerable to exploitation in the form of a specifically engineered set of trades or bets guaranteed to exploit this time inconsistency, known as a "Dutch book" **(Callender, 2021).** "An optimal schedule of consumption is time consistent if it is still optimal when reconsidered at some later time **(Callender, 2021).**" This exploitability is why exponential discounted utility theory became normative **(Callender, 2021).** However, not all variations in preference indicate violations of consistency. Specifically, we can distinguish three similar, yet distinct properties of intertemporal choice - **stationarity, consistency** and **invariance.**

An agent abides by the axiom of stationarity if the ranking of two choices depends only on the values of the choices at their respective times of consumption and the delay between those choices. If we prefer 110 dollars in 8 days over 100 dollars in a week, to abide by the stationarity axiom, we must also prefer 110 dollars tomorrow over 100 dollars now.

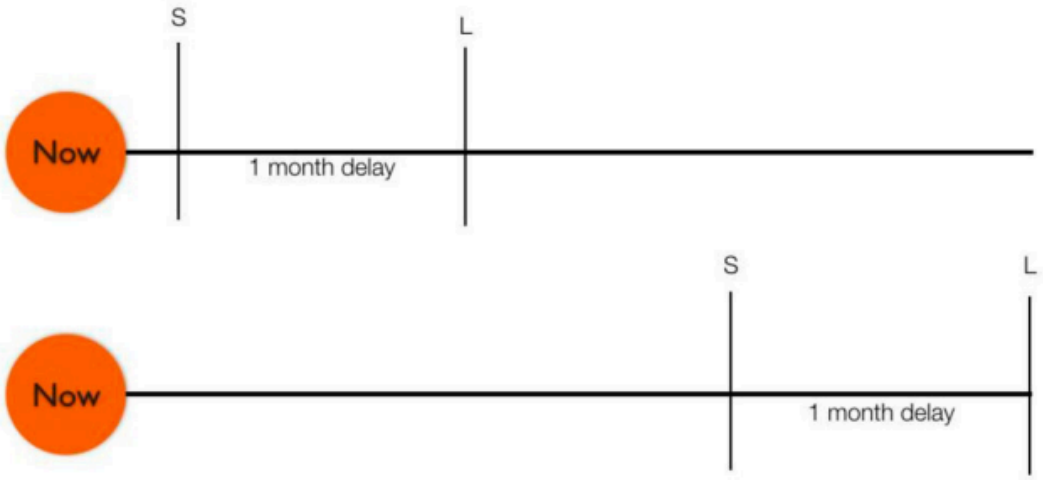

Figure 2.1 (Callender, 2021): "Stationarity: Let the horizontal line represent a timeline, the dot the evaluation point, and S a small reward and L a large reward. A set of preferences that is indifferent between the top and bottom situations satisfies Stationarity."

However, violating stationarity isn't necessarily exploitable, as there isn't an apparent preference reversal – we simply prefer, at the same time, A over B when they are close and B over A when they are far, but this is an evaluation of two different

outcomes, and can have many explanations. For a true preference reversal, an agent must violate **consistency**. An agent violates consistency, when he changes his preferred choice over time without any additional changes other than the passage of time. We would violate consistency, if we preferred 110 dollars in 8 days, but when given the option to reconsider after a week, we would take the immediate 100 dollars instead of waiting the extra day for 110 dollars, which we set out to get beforehand.

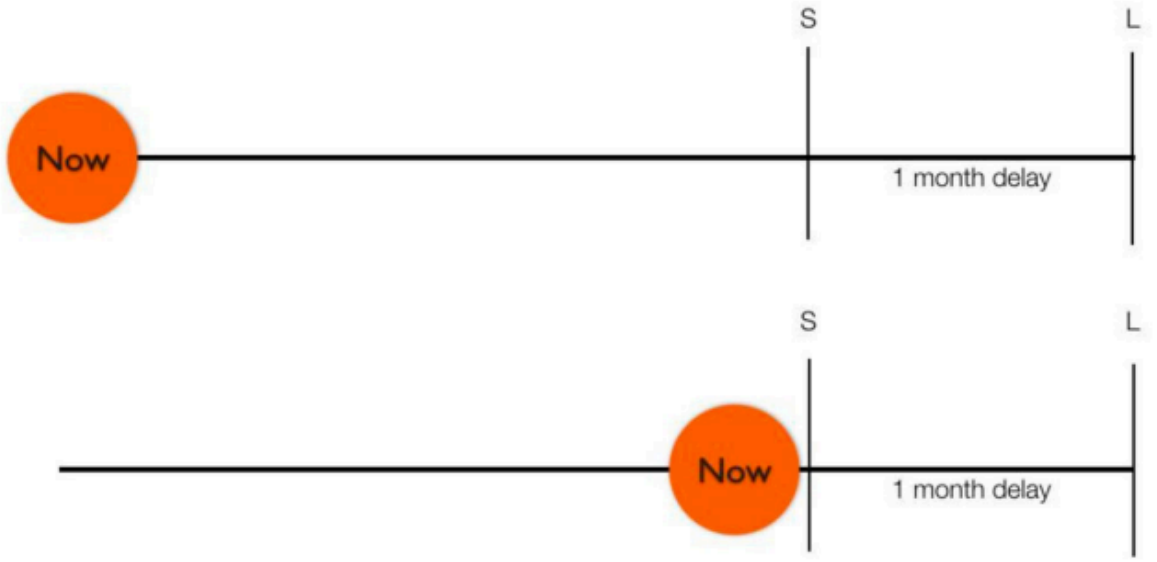

Figure 2.2 (Callender, 2021): "Consistency: Let the horizontal line represent a timeline, the dot the evaluation point, and S a small reward and L a large reward. A set of preferences that is indifferent between the top and bottom situations satisfies Consistency."

However, not all preference reversals after time has passed can be attributed to the violation of consistency. Changes in preferences can also be attributed to violations of **invariance**. Such violations occur when preferences change over time – one time, we may prefer 110 dollars both 8 days away and 1 day away, but then, when asked after some time has passed, we may prefer 100 dollars immediately or after a week. This behavior would violate invariance.

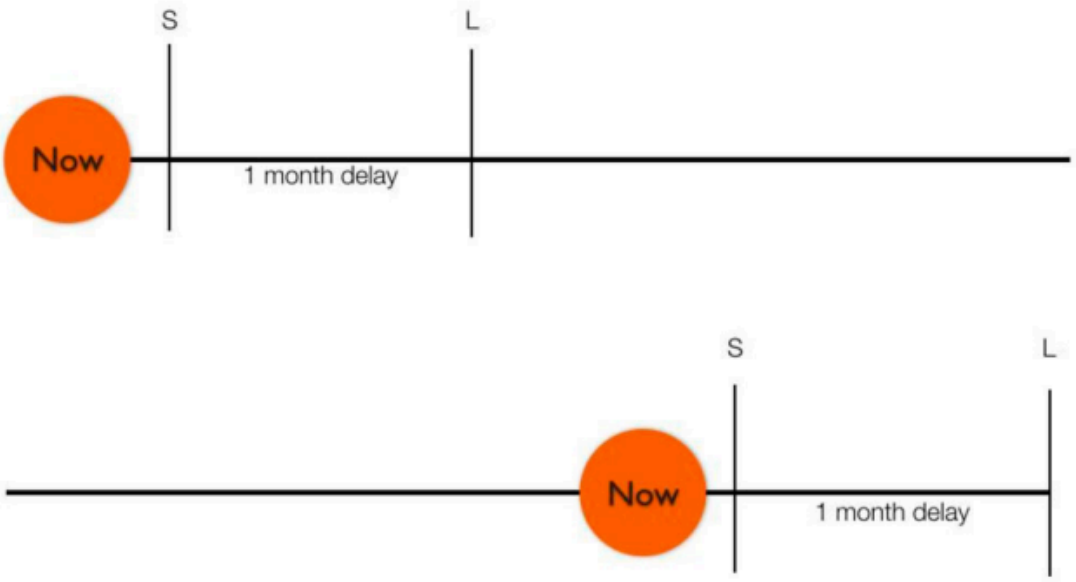

Figure 2.3 (Callender, 2021): "Invariance: Let the horizontal line represent a timeline, the dot the evaluation point, and S a small reward and L a large reward. A set of preferences that is indifferent between the top and bottom situations satisfies Invariance."

Callender's main criticism of contemporary research on time preferences was that numerous studies claimed violation of consistency, while not accounting for invariance **(Callender, 2021).**

While **Samuelson (1937)** hinted at the limited ability of exponential discounted utility to describe actual human behavior, the first to seriously investigate the topic was **Robert H. Strotz (1955)**, who noted that consumers display utility myopia, and are dynamically inconsistent, i.e., violate consistency.

We can clearly see the appeal of exponentially discounted utility, which contributed to it becoming a normative standard for future discounting, in Figure 2.4. An agent discounting all future utility exponentially with the same rate will have time-consistent preferences.

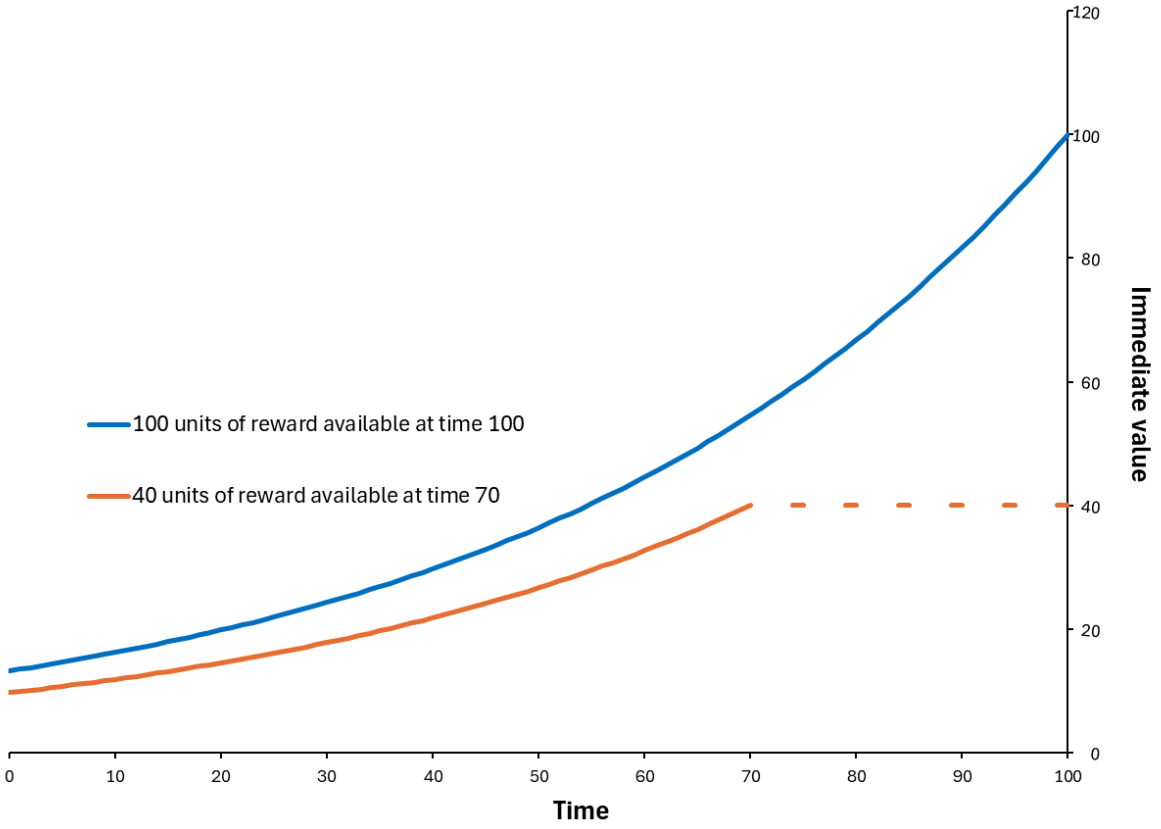

Figure 2.4: Exponential discounting of two rewards with different sizes and times of consumption. At each point, the discounted value of the larger reward is bigger than the discounted (or immediate at time 70) value of the smaller reward. At no point does exponential discounting violate consistency.

In some cases, time-inconsistent preference reversals can be perfectly rational. "According to most choice theories, rational decisions should be exclusively prospective [relating to the future], and not retrospective [relating to the past] **(Kalenscher & Pennartz, 2008, p. 290)**." Interestingly, it has been observed that humans and other animals not only display dynamic inconsistencies, i.e., time-inconsistent behaviors violating the axiom of consistency, despite lack of new information that would warrant it, they also do the exact opposite by maintaining their choices despite the emergence of information that would warrant changing their preferences **(Kalenscher & Pennartz, 2008)**. This is known as the sunk cost effect or sunk cost fallacy, as it is commonly associated with the aversion of the loss of resources already spent pursuing the choice that should now be abandoned. Several mechanisms of motivation for such behavior have been proposed, "...including hesitation to waste resources (Arkes and Ayton, 1999), loss

aversion (Schaubroeck and Davis, 1994), desire to complete a job (Moon, 2001) or project-based mental accounting (Heath, 1995)**(Kalenscher & Pennartz, 2008, p. 290).**"

Matching law

Now, after introducing normative economic theories of intertemporal choice, we will investigate their descriptive validity by reviewing empirical findings on intertemporal choice. We therefore shift focus towards behavioral research. While most original empirical behavioral research has been done on animals, "virtually all of the empirical generalizations found regarding animal choice have also been found to hold regarding human choice (e.g. McDowell 1988; Rachlin 1987) **(Vuchinich & Heather, 2003, p. 2).**"

In 1961, Richard Herrnstein conducted an experiment, which would become the basis for one of the original behavioral theories of choice – the matching law **(Herrnstein, 1961, 1970)**. In this experiment, Herrnstein let pigeons peck one of two keys independent of each other, each of which provided the same amount of reinforcement (food), but with different variable intervals. This means that following a peck of a key which resulted in obtaining reinforcement, some time had to pass, before pecking the same key resulted in more reinforcement. Keys being independent on each other meant that pecking one key had no effect on the other key. This time was partially randomized around set means such that the total available reinforcement from the two keys stayed the same. What Herrnstein found is that the relative rate of reinforcement closely matches the relative rate of pecking the key, hence the name matching law. If one key has an average interval of one minute and the other has an interval of two minutes, matching law predicts that the one-minute-interval key will be pecked twice as often. This can be represented very simply as an equation:

$$\frac{B_1}{B_2} = \frac{FR_1}{FR_2}$$

In this equation, B_1 and B_2 represent behavioral allocation, or individual responses allocated to options 1 and 2. FR_1 and FR_2 then denote respective frequencies of reinforcements for options 1 and 2. **Catania (1963)** subsequently found that the same law

holds when the means of variable intervals are the same but amounts of reinforcement differ:

$$\frac{B_1}{B_2} = \frac{AR_1}{AR_2}$$

Where AR_1 and AR_2 represent amounts of reinforcement from options 1 and 2.

For us, the most important observation was that behavioral allocation showed to be inversely proportional to the relative delay of reinforcement **(Chung & Herrnstein, 1967)**:

$$\frac{B_1}{B_2} = \frac{DR_2}{DR_1}$$

Where DR_1 and DR_2 represent the delays of receipt of reinforcement after choosing options 1 or 2. This goes against exponentially discounted utility theory, because relative delay of reinforcement isn't time-consistent, therefore behavioral allocation also isn't time-consistent – according to the matching law, a key that provides reward after a delay of 5 seconds on average will be pressed twice as much as a key that provides reward after 10 seconds on average, yet a key that provides a reward after 10 seconds on average will be pressed only one and a half times as much as a key that provides reward after 15 seconds on average.

The matching law, unlike exponentially discounted utility theory, predicts the occurrence of preference reversals in favor of sooner rewards, as they move closer in time, "even though nothing differs between the choice points except for the temporal distance from reinforcement **(Vuchinich & Heather, 2003, p. 9)**". Vuchinich and Heather then provide a hypothetical example of such preference reversal, shown on Figure 3.1. When calculating the values strictly according to the matching law, that is, divide the reward amount by the delay, we see that at choice point Y, that is farther from both rewards, larger, but later reward is preferred. At choice point X, however, sooner, but smaller reward is preferred. Corresponding equation to calculate reward's value V as a function of amount A and delay t, according to the original matching law would then be:

$$V(A, t) = \frac{A}{t}$$

This equation is therefore a form of temporal discounting function, and the function has a hyperbolic shape.

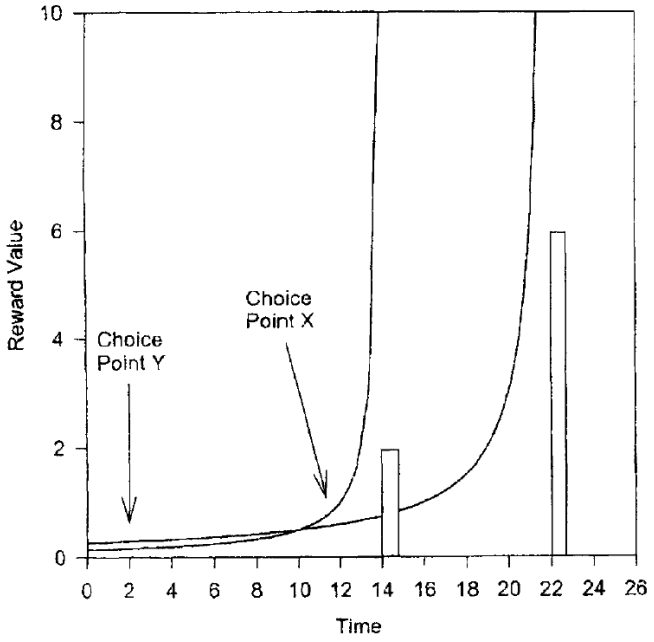

Figure 3.1 (Vuchinich and Heather, 2003, p. 10): "Intertemporal choice between two seconds of food available at time 14 and 6 seconds of food available at time 22."

Hyperbolic discounting

The simple equation that immediate value equals reward amount divided by delay comes with two issues. Firstly, as the delay approaches zero (immediate reward), the value approaches infinity. This means that however small, immediately available reward would always be preferred over delayed reward, however big. Secondly, it implies that all individuals discount all delayed rewards at the same rate. As such, it cannot be used as a true predictor of behavior, because both assumptions contradict empirical evidence **(Ainslie and Monterosso, 2003).**

It was, however, the first step to the finding that the rate of temporal discounting is not constant with each unit of delay, instead, it is dependent on the proportion to overall delay. This is represented as a temporal discount function with delay in the denominator **(Ainslie and Monterosso, 2003).** The simplest and most commonly referred to formula is **Mazur's (1987):**

$$V(A,t) = \frac{A}{1+kt}$$

where V is current value, A is the undiscounted reward value, t is delay, and k is a constant which describes the degree of myopia, impatience, or propensity to overvalue immediate rewards.

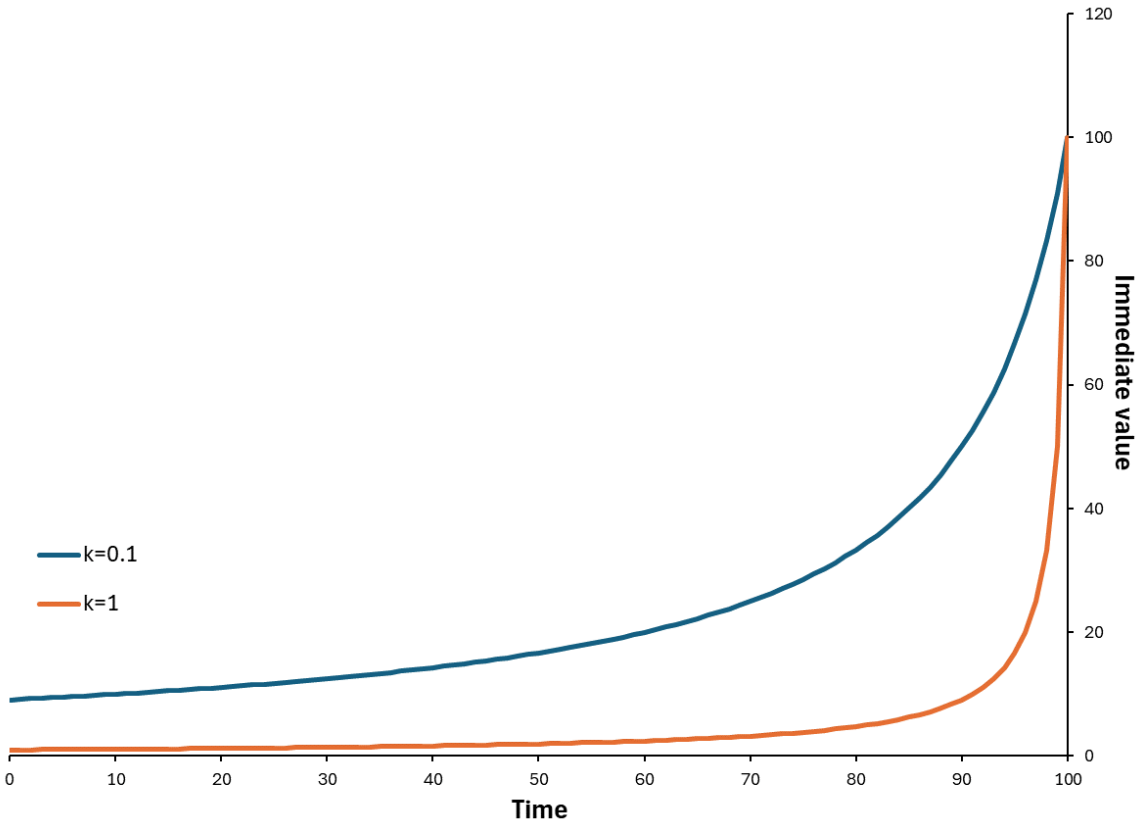

Figure 3.2: Reward worth 100 units at the point of consumption available at time 100 hyperbolically discounted using Mazur's (1987) formula for k=0.1 (top curve) and for k=1 (bottom curve)

Hyperbolic vs. exponential discounting

Into the same example as in Figure 3.2, I also plotted an exponential function to showcase the differences between hyperbolic and exponential discounting. The exponential discount rate was chosen arbitrarily as 5% (δ=0.95) per unit of delay.

While this isn't a universal rule for all pairs of exponential functions and hyperbolic functions calculated by Mazur's formula, for qualitatively similar (i.e., can be fitted to the same data), two phenomena occur:

For delays in the short to middle ranges, hyperbolic discounters discount future utility more steeply than similar exponential discounters. Therefore, in these ranges, delaying the reward has a bigger negative impact on its value in the case of hyperbolic discounter. For delays in the long ranges, exponentially discounted utility of rewards which are very far away becomes negligible. Comparatively, hyperbolic discount functions have "long tails", i.e. very far away, hyperbolically discounted rewards aren't discounted steeply at all. Put simply, "the more delayed an option is, the less discounted it is **(Ainslie, 2001, p. 44)**"

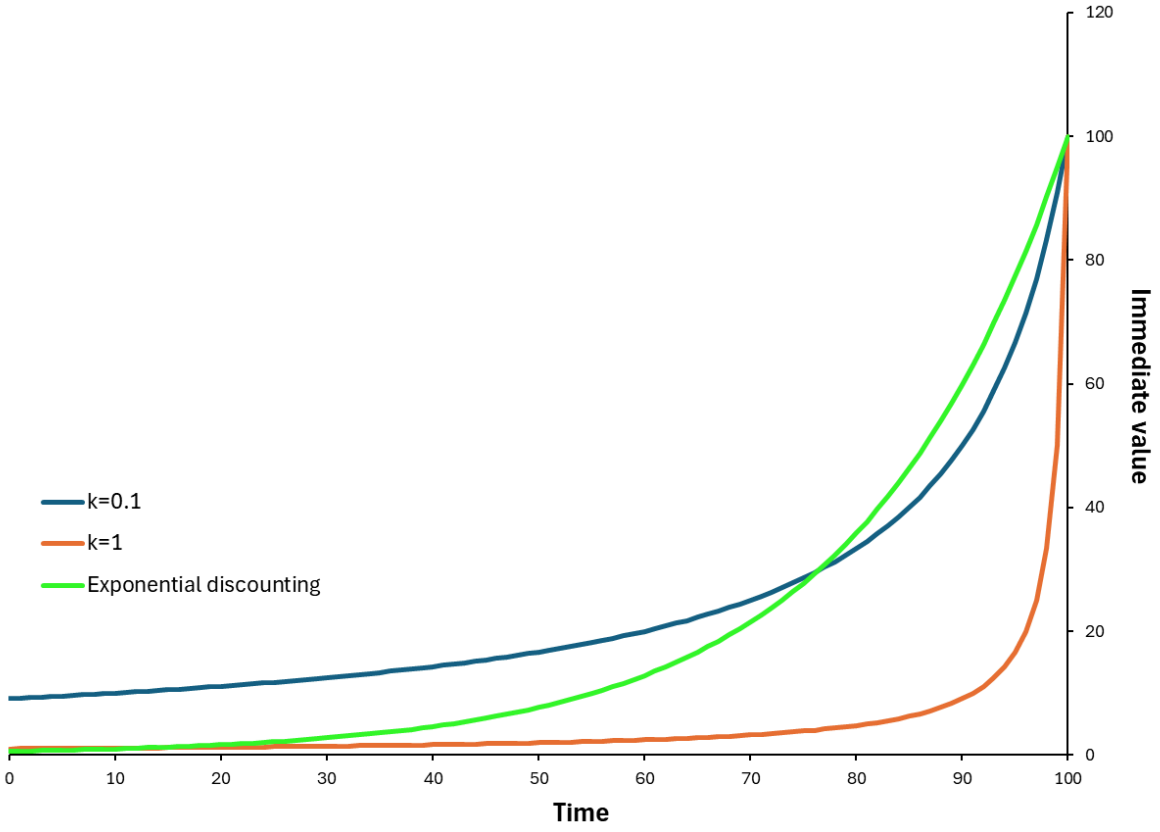

Figure 3.3: When it is far away, hyperbolic discounters discount future reward less steeply than exponential discounters. When it is near the point of obtainment, hyperbolic discounters discount future reward more steeply than exponential discounters.

Of course, we should somehow define what are short to medium ranges of delay, and what are long ranges of delay, especially since the examples use arbitrary units. Defining an universal cut-off point where a hyperbolic discounter starts to discount utility less steeply than an exponential one may not be possible due to significant differences in

discounting based on the experiment design, the individual discounter, and the goods used to elicit the discount rate. Research on animals, which used reinforcers such as food, juice or aversive noise and which has been the first step to identify steep time-inconsistent discounting of immediate rewards has been done in the ranges of seconds to minutes **(Chung and Herrnstein, 1967; Ainslie and Monterosso, 2003)**. Research on the longest-range planning done by individuals, such as choices on preserving the environment or saving money for grandchildren can have benefits tens to hundreds of years from now showed virtually no discounting at all **(Ainslie, 2001, p. 44)**.

Evidence points to the conclusion that **hyperbolic discount functions better describe animal (including human) future utility discounting compared to exponential discount functions.** They do so by predicting proportionally higher rates of discounting of temporally close rewards compared to those temporally far away, which has been consistent with research over different rewards, time horizons and species. Most importantly, they predict the occurrence of **preference reversals**, so that a shorter reward, that is considered suboptimal from a distance, can be preferred when immediately available (Figure 3.4).

An agent discounting like in the figure 3.4, if given the option to choose from a distance, would prefer the larger, later reward. However, if given the choice, when the smaller reward is near (less than 10 time units), he would prefer the smaller, sooner reward. This dynamic inconsistency is especially problematic for choices which are mutually exclusive and available for the whole duration (the agent cannot precommit, i.e., remove the smaller, sooner reward ahead of time).

Such choices occur often in real life. Imagine a scenario where we want to lose weight and are invited to dinner at a restaurant. The day before, we plan to skip on dessert (the smaller reward), so that we do not sabotage our diet efforts (in this case weight loss is the larger, later reward). However, at dinner, once we have finished our main course and the waiter asks us if we would like dessert, we often give in to the temptation and despite our previous plans to forego it, we order a piece of chocolate cake.

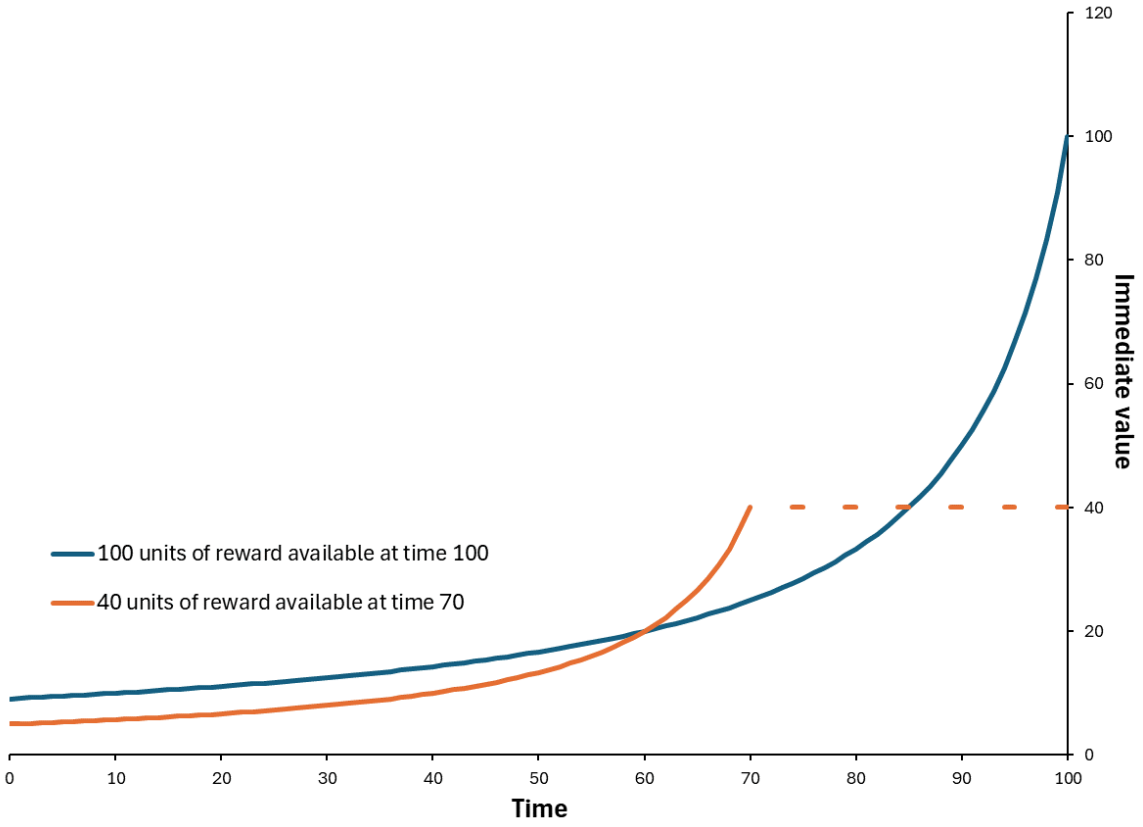

Figure 3.4: Preference reversal predicted by hyperbolic discounting, k=0.1 (which is fairly conservative and not very steep). From a distance (up until time 60), the larger, later reward is preferred over smaller, sooner reward. At time 60, when the reward of 40 units is 10 units of time away, and the reward of 100 units is 40 units of time away, their discounted value of both is 20 reward units. After that, the discounted value of the smaller reward is larger than the discounted value of the larger, later reward.

There also emerged other mathematical functions set out to model temporal discounting consistent with empirical evidence, such as **quasi-hyperbolic (Laibson, 1997).** Quasi-hyperbolic discounting assumes exponential discount rates for all delayed utility, with added value for immediately available choices known as immediacy premium or present bias, denoted by β, a factor by which all delayed choices are devalued:

$$V(A,t) = A \qquad \text{if } t = 0$$

$$V(A,t) = A \times \beta\delta^{t} \quad \text{if } t > 0$$

with β usually being defined on the interval [0; 1]. Value β of 1 implies no present bias and discounting mirrors **Samuelson's (1937)** exponential discounting. Value β of 0 would

imply infinitely steep discounting, where all non-immediate rewards have no value to the agent. The appeal of quasi-hyperbolic discounting function is its computational simplicity and analytical tractability while retaining the qualitative properties of hyperbolic models **(Laibson, 1997)** coupled with the ability to easily discern between present bias and long-term discount rates using the two parameters.

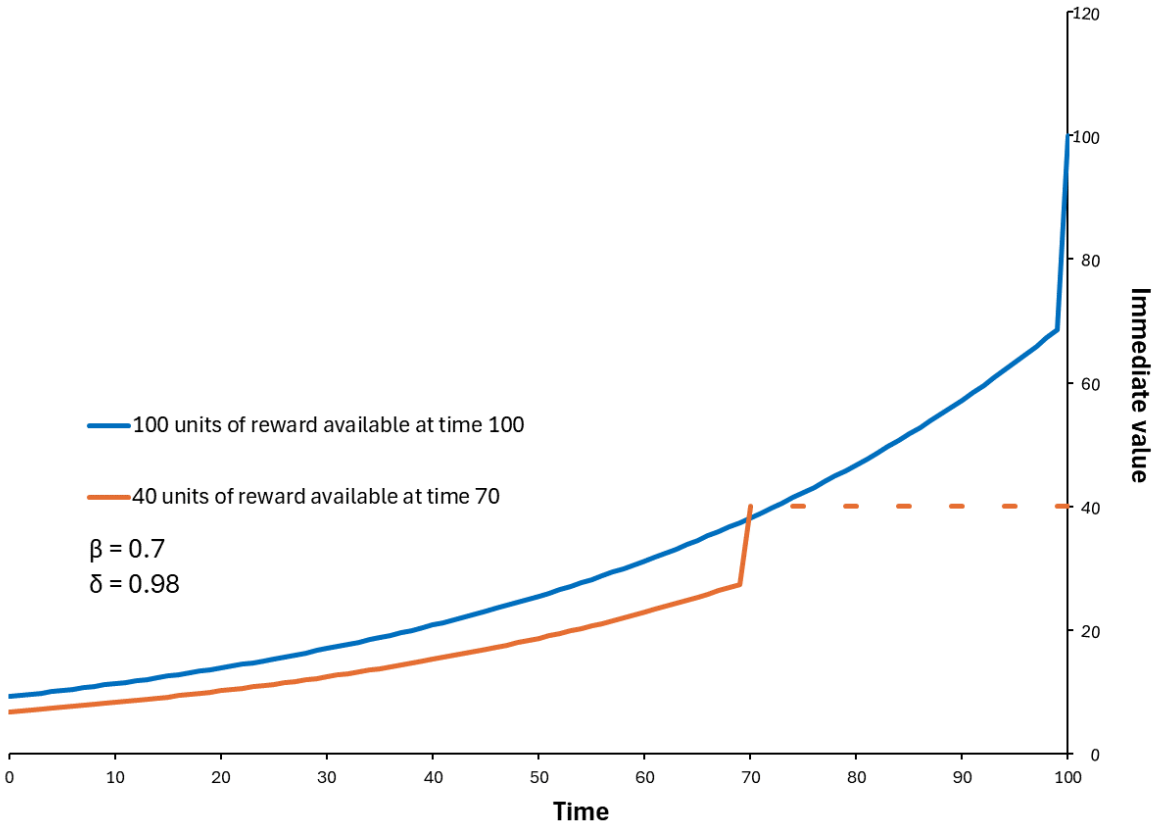

Figure 3.5: Dynamic preference reversal shown using quasi-hyperbolic discounting, with immediacy premium parameter $\beta = 0.7$ and exponential discounting parameter $\delta = 0.98$.

Other, more sophisticated models have been proposed, which seek to better address some of the limitations of the original models, such as the unintuitive interpretation of the values of k in **Mazur's (1987)** formula and subsequent issues with interpreting correlations, or the occurrence of non-hyperbolic violations of time consistency **(Urminsky & Zauberman, 2016)**. Those seeking to find suitable models of temporal discount function for their specific use case should refer to other sources **(see, e.g., Urminsky & Zauberman, 2016 for sources)**, as detailed analysis is outside the scope of this thesis.

The role of discount functions in intertemporal choice research

There are several ways to understand these discount functions with regards to actual decision making processes. One would be to assume that there exists a unitary process inside the brain in which precise or approximate calculation of discounted utility is made according to a set discount function. Such a mechanistic view, if correct, would simplify our understanding of the topic and would increase the predicting power of economic models utilizing these discount functions. This view, however, contradicts the contemporary understanding of decision making, and I would caution against it.

Another way is to understand decision making, including intertemporal choice, as a set of processes which seem unitary to the outside observer, but can be broken down further as a cooperation and competition between diverse centers and psychological processes. Using this view, discount functions estimated by observing choices done by these sets of processes over time should (ideally) approximate optimal discount functions for the environment in which these sets of processes evolved, but different processes could influence different time horizons. For example, visceral, evolutionarily older processes could exert large influence over immediately available choices, while higher reasoning and cognition could cause ourselves to consider the medium and long-term future. A sense of responsibility for the environment or the wellbeing of future generations could then transcend even a relatively long-term horizon and be responsible for the long tails of the discount function. Through cooperation, competition and activation at different time horizons, these processes would then shape the discount curve using their pushing (e.g., willpower and deliberation) and pulling (e.g., urges and impulses) forces.

Both views face two challenges. Firstly, it is crucial to determine the optimal discount function which pertains to an environment. Although neoclassical economists considered this issue solved by proving the normativity of exponential discounting **(Samuelson, 1937; Strotz, 1955)**, **Peter Sozou (1998)** showed that while exponential discounting is advantageous in environments with a known hazard rate, **uncertain hazard rate implies that an agent ought to discount hyperbolically**. By hazard rate, we mean the risk of death of the agent. This threatens the foundational principles set forth by expected utility theory, because if we agree that in our real world environment, there is

some kind of uncertain hazard rate, the very basic notion that exponential discounting is rational is threatened.

Another challenge is to then determine the actual function which is supposedly either computed or composed. Only after determining both optimal and actual discount functions can they be compared to identify systemic departures from optimal choice, which would give grounds for the discussion of irrationality.

Instead of focusing on the precise mathematical properties of discount functions and analyzing all possible discounting models that have been proposed, I chose to present only some of the most well known, i.e., Mazur's hyperbolic formula, and Laibson's quasi-hyperbolic discounting. Their role in this thesis is mainly to visualize and help better comprehend the diminishing rates of discounting with increases in delay and myopic preference reversals when smaller, sooner rewards are near, instead of claiming any knowledge about the precise mathematical formulation of a function which would best describe human intertemporal choice. This decision risks losing valuable information about phenomena which are function-specific and very sensitive to the shape of the discount function. For example, **Kurth-Nelson and Redish (2010)** found that a relatively small change in the approximation of hyperbolic discount function, which influenced the "long tail" of the function, had an influence on whether the model precommits to the larger reward or not, i.e., whether the model limits its option to choose the smaller, sooner reward in the future, when it will be immediately available. However, as will be shown later, there are many distinct mechanisms which could be responsible for myopic choices and dynamic inconsistencies, many of which can but do not have to exist together, some of which can be relevant to certain outcomes but irrelevant to others, which could mean that there may not even exist a unitary utility discount function within an individual, let alone across individuals. It is therefore hardly surprising that it was shown that in humans, discounting is very context and factor-specific, and discount curves derived from observing human choices are variable, unstable, poorly correlated between different goods and often unrelated to subsequent life choices, greatly undermining any notion regarding their predictive power **(Ainslie and Monterosso, 2003)**. As **Roelofsma and Read (2000, pp. 171–172)** have put it,

> There have been many recent articles attempting to specify the form of the individual discount function. The typical finding is that hyperbolic discounting fits observed behavior better than exponential discounting (e.g. Benzion, Rapaport and Yagil, 1986; Kirby, 1997; Overton and MacFadyen, 1998). There is considerable doubt, however, whether the psychological processes underlying this behavior actually draw on a personal discount function [...]
> Decision makers appear to have as many discount rates as choice situations into which they can be placed. Moreover, different measures of discount rates are either uncorrelated, or are correlated weakly or idiosyncratically (e.g. Chapman and Coups, 1998; Chesson and Viscusi, 2000; Vuchinich and Simpson, 1998).

And even despite the finding that hyperbolic discounting does a better job of explaining intertemporal choice, "researchers have demonstrated patterns of choice that seem anomalous even from the framework of hyperbolic discounting **(Ainslie and Monterosso, 2003, p. 40)**." For example, **Ainslie and Monterosso (2003)** note that addicts sometimes plan their consumption in advance, undermining the notion that it is dependent only on close temporal proximity, and **Thaler and Sunstein (2008, p. 107)** provides several examples of people not taking advantage of opportunities for considerable sums of money for low effort, such as not filling out a simple form for a retirement account to which only the employer contributes. Therefore, assigning people hyperbolic discount functions while still treating them as an unbiased "homo economicus" may move us nowhere closer to understanding their observed economic behavior or the mechanisms behind it.

Intertemporal research limitations and challenges

Consider, for example, the most common good upon which discounting is investigated – money. Common designs of experiments attempting to determine properties of temporal discounting are either asking people to determine equivalent value after delay (e.g., How many dollars would you have to receive in 7 days to prefer them over 10 dollars now?), or letting them choose between pairs of choices (e.g., Would you rather receive 5 dollars now, or 10 dollars in 7 days?) **(Urminsky and Zauberman, 2016)**. There are numerous limitations to empirical research done this way. For simplicity, in temporal discounting research, it is often assumed that utility from a good is received only at the point of consumption. Only some researchers incorporate any notion of anticipation utility in these simplified models **(e.g., Banerjee and Mullainathan, 2010)**. Yet, in real life, even short-term consumable goods such as food have a spread out window of giving us utility. For example, we can derive utility from anticipating the food, eating the food, being satiated for some time after eating the food, and even from remembering eating the food **(Callender, 2021)**. Therefore, research that simplifies the model by constraining utility from a reward into a single point in time has to address this difference.

Using money is especially problematic, because it is a medium of exchange, and its consummatory period is unspecified **(Ainslie and Monterosso, 2003)**. A well-off individual can readily substitute the smaller, sooner monetary reward with their own money, and can only gain by waiting for the larger reward. On the other hand, an extremely poor individual can find themself in a situation where they must choose the immediate monetary reward to survive. Despite being offered the same choice in terms of value, the (negative) utilities stemming from postponing the monetary rewards are incomparable. It has been observed that discount rates of money in very poor households can be in the ranges of thousands of percent per year **(Tanaka et al., 2010)**, while for households that are not financially struggling, it is much less, often assumed to be in the range of tens of percents per year **(Andersen et al., 2008)**, which gives us some insight into how differently well-off individuals discount money, much less, however, into how differently well-off individuals discount utility.

Similar discount rates have been observed for hypothetical, probabilistically paid and paid money **(Urminsky & Zauberman, 2016)**. However, each implementation has its own challenges. For hypothetical monetary rewards, the ability to imagine hypothetical outcomes can influence the extent to which hypothetical choices match real choices **(Keidel et al., 2021)**. For real or probabilistically paid monetary rewards, seemingly steep discounting can be a result of a lack of trust towards the researcher to actually pay out the delayed reward **(Ubfal, 2016)**.

Members of a population that is studied may be also inclined to please the researcher and act in a way that supports their hypothesis, for example, drug dependent individuals could deliberately choose smaller, sooner rewards, since for the duration of the experiment, they feel defined by their drug addiction, because it is the defining distinctive characteristic in the experiment **(Ainslie and Monterosso, 2003)**.

Research on discounting was also done with many other goods and outcomes besides money, such as various foods **(Ubfal, 2016)**, cigarettes **(Story et al., 2014)**, alcohol **(Story et al., 2014)**, illicit substances such as heroin or cocaine **(heroin: Story et al., 2014; cocaine: Bickel et al., 2011)**, but also health outcomes **(Story et al., 2014; Keidel et al., 2021)**, freedom **(Petry, 2003)**, and even having to listen to aversive noise, or looking at pictures of celebrities **(Ainslie and Monteresso, 2003)**. It was found that overall, there are many factors which influence discount rates and discount function shapes, which must be controlled for to ensure unambiguity of research results. We will review these factors in the next chapter.

Special consideration must also be given to any claims of causality. Most research on intertemporal choice has been done by finding correlations. While in some cases, it could be intuitive to think that, for example, steep discounting causes addiction, causal relationships are still unclear. It may as well be the case that addiction makes changes in the brain which manifest as steep discounting, or that something else completely causes both addiction and steep discounting **(Ainslie & Monterosso, 2003)**.

Factors influencing temporal discounting

Individual differences

One of the most important factors that temporal discounting research focuses on is individual differences. It has been found that there is considerable variability in the rate of temporal discounting not just across different rewards, but also across different individuals **(Keidel et al., 2021).** One objective is to estimate the discount rates of individuals to identify phenomena associated with different discount rates. Another objective is to find patterns within the populations, such as similar rates of discounting across certain subpopulations (e.g., smokers, alcoholics, young adults, religious people, poor individuals). Identifying such patterns could help shed light into phenomena observed within these individuals and subpopulations despite the unclear causal relationship between steeper discounting and certain pathologies **(Keidel et al, 2021).**

Age

Age seems to be a determining factor of temporal discounting, although the exact relationship is far from being clearly understood and defined, with different studies showing different results **(Keidel et al., 2021).** It seems though that discount rates are higher during childhood and adolescence and decline over time, but also sometimes increase back in older adults (51–75 years), possibly due to cognitive impairment **(Keidel et al., 2021).**

Gender

Studies on effects of gender have also not provided a unitary answer, but a meta-analysis argued in favor of men having higher discount rates than women **(Gaillard et al., 2020).**

Education level

Educational attainment also seems to be linked with temporal discounting, with higher education suggesting lower discounting, as was also the case for higher **income**, **wealth** or

home ownership (Keidel et al., 2021), suggesting a correlation between patience and professional success. Again, it is hard to determine a causal relationship.

Impulsivity

Differences in discounting have also been associated with different personality traits, most notably impulsivity. While not the case for all traits considered impulsigenic, lack of premeditation, along with behaviors considered impulsive such as substance use and gambling have been linked to steeper discounting **(Keidel et al., 2021).** Other personality traits that have been linked to steeper discounting, although only with small effect, have been higher extraversion, lower openness and conscientiousness, and higher neuroticism **(Mahalingam et al., 2014).** From the same domain, certain psychiatric disorders were also linked with steeper discounting, such as ADHD, schizophrenia, and depression **(Keidel et al., 2021).**

Religion

Identifying oneself as a religious person has been linked to lower rates of drug dependence, which also suggests shallower discounting **(Ainslie & Monterosso, 2003).**

Mental imagining

An interesting relationship is between temporal discounting and mental imagery ("the ability to form and experience mental representations of stimuli without actually perceiving them; **Keidel et al., 2021, p. 7**"). From one perspective, greater ability to imagine the future should be linked with choosing larger, later rewards due to their higher salience, and some studies confirm just that **(Keidel et al., 2021).** However, the same ability could also allow us to better discern small differences between large delays (e.g. 100 vs 101 days), making us prefer the smaller, sooner reward even from a distance. In this case, the dynamic inconsistency stems not from an inability to resist immediate reward, but from our imperfect ability to imagine large delays **(Kalenscher & Pennartz, 2008).** This is also supported by empirical evidence **(e.g. Parhasarathi et al., 2017).** While having mixed evidence about the effect on the steepness of discounting, it does

raise the question of whether the level of the ability to imagine the future has any effect on dynamic inconsistency of temporal choice.

Intelligence

Research was also done in relation to the effect of intelligence on temporal discounting, which found that higher intelligence is correlated with shallower discounting, however, the exact nature of how intelligence is connected to discount rates is not fully understood, since intelligence is connected to other relevant traits, such as the ability to imagine future outcomes, or cognitive reflection **(Keidel et al., 2021; Urminsky & Zauberman, 2016)**. Urminsky and Zauberman mention a meta-analysis by **Shamosh and Gray (2008)**, which found a significant negative relationship between intelligence and discounting (i.e., higher intelligence means lower discounting).

Heritability

It has been found that discount rates are moderately but significantly heritable **(Keidel et al., 2021)**.

Substance use and addictive behaviors

The link between substance use and steeper discounting is well established. While discount rates for money seem poorly correlated with discount rates for health outcomes, the relationship between monetary discount rates and substance use are more consistent **(Story et al., 2014)**. Findings from research on smokers suggests that monetary discounting is related to the level of current dependence – smokers discount money more than non-smokers, more frequent smokers discount money more than less frequent smokers, administering nicotine to non-smokers didn't increase their discount rates, smokers who were currently craving a cigarette discounted money more steeply than smokers who had one recently, and most importantly, it's been shown that ex-smokers have discount rates similar to non-smokers **(Story et al., 2014)**. However, there is also evidence that high monetary discount rates increase both the chance of future smoking adoption, as well as of relapse, and **Story et al. (2014, p. 5)** therefore conclude that "relationships between discounting and smoking behavior are subject to both state- and

trait-based influences." Alcohol and certain other addictive substances seem to have a similar relationship to monetary discounting, with dependence, more frequent use, as well as state of withdrawal being associated with steeper discounting **(Story et al., 2014).**

It is not surprising, then, that considerable effort has been made to identify both behaviors associated with steeper or shallower discounting, or the differences in discounting across individuals and subpopulations. Such research can help shed light into certain aspects of personality forming, and most importantly, could be the first step to the development of tailored solutions to long-term utility maximization for different phenotypes of discounters. However, results of research on isolated traits and variables must be interpreted carefully, especially when trying to identify causal relationships **(Keidel et al., 2021).** Also, connecting a subpopulation on one trait can disconnect it on another, intercorrelated trait **(Ainslie and Monterosso, 2003).**

Qualitative and quantitative differences of rewards

Magnitude

Absolute sizes of rewards that are compared play a role in the steepness of their discounting **(Loewenstein & Prelec, 1992; Ainslie and Monterosso, 2003). Loewenstein and Prelec (1992)** call this The absolute magnitude effect. They mention **Thaler's (1981)** research, which found that "subjects who were on average indifferent between receiving $15 immediately and $60 in a year, were also indifferent between an immediate $250 and $350 in a year, as well as between $3000 now and $4000 in a year. Similar results were obtained by Holcomb and Nelson (1989) with real money outcomes **(Loewenstein & Prelec, 1992, p. 575).**"

Gain-loss asymmetry

There also seems to be a sign effect **(Kalenscher & Pennartz, 2008)**, or an asymmetry between gains and losses, with losses being discounted less steeply than gains **(Loewenstein & Prelec, 1992).** This finding largely mirrors prospect theory, in that a

delayed loss looms larger than a delayed gain of the same amount, just as a risk of a loss looms larger than a chance of a gain of the same value **(Kahneman & Tversky, 1979).**

Framing

People's intertemporal choices are also dependent on the way in which they are presented, and are thought to form a reference point (see chapter Objectivity of time).

Order effect

In the case of a series of choices or a series of rewards, we can also see an order effect **(Urminsky & Zauberman, 2016).** Series of rewards which are in ascending order (i.e., from least valuable to most valuable) are preferable over a series of rewards which are in descending order **(Kalenscher & Pennartz, 2008)**, and are discounted less steeply **(Urminsky & Zauberman, 2016).**

Good-specific discount rates

Contrary to the normative assumptions of neoclassical economists about the irrelevance of the source of utility, it has been repeatedly shown that people discount different goods and outcomes differently, which is another possible explanation for time-inconsistent choices. For example, **Ubfal (2016)** found that on average, people discount certain staple foods (meat, sugar, plantains) more steeply than money, and some goods, such as shoes or school supplies, less steeply than money. Moreover, research suggests many other goods and outcomes are discounted more steeply than money, such as time **(Urminsky & Zauberman, 2016)**, addictive substances such as cigarettes, alcohol or heroin **(Story et al., 2014)**. For discounting health outcomes, results have varied, with some authors suggesting steeper discounting than that of money, and others suggesting shallower **(Urminsky & Zauberman, 2016).** There are several challenges when comparing the discounting of different rewards, such as finding equal values, so that differences in discounting cannot be attributed to the magnitude effect, or differences in item-specific utility functions, such as variations in diminishing marginal utility **(Urminsky & Zauberman, 2016).** It is also possible that the brain uses different heuristics when evaluating the discounted value of qualitatively different rewards **(Ainslie & Monterosso,**

2003). Despite these challenges when comparing discounting of different rewards, I will review some qualities of different goods and outcomes which influence their future discounting.

Salience

Salience is an important factor when evaluating the future value of a good. Goods and outcomes which are salient, i.e., prominent, vividly imagined, are more likely to be considered. While it is not entirely clear whether salience and steepness of discounting are interconnected, or are two distinct mechanisms, it was shown that manipulating salience of future outcomes influences discounting **(Urminsky & Zauberman, 2016).** Students who were merely reminded to consider the long-term effects of their choices displayed shallower discounting that when they were not **(Ainslie & Monterosso, 2003).** A highly relevant factor to salience is the **objective evaluation** of an outcome. One of the factors why people seem to discount time more steeply than money is that differences in delay are harder to evaluate than differences in monetary value, e.g., isolated duration of three months can feel longer than when it is a part of a year-long interval **(Urminsky & Zauberman, 2016).**

Affective activation

Affective activation of a reward is another quality of a good which influences the way in which it is discounted. By affective activation, it is meant the effect that it has on our emotions, urges, instincts, unconscious processes and all that many behavioral economists call the "hot" system, especially when the good is readily available, or an appetite for it has been created by external or internal cues. While all goods, including delayed ones, most likely elicit some kind of affective activation **(Ainslie, 2001, p. 25)**, I will mostly focus on so-called tempting goods. There are different ways to define tempting goods, for example **Banerjee et al. (2015)** define tempting goods as goods by which people feel tempted and wish to consume less than what they currently consume. For the purpose of this thesis, I consider the most important quality of a tempting good that it provides hedonic utility (i.e., immediate pleasure) when consumed, and a cost that is incurred at a later time. Examples of such goods (or behaviors) are affectively rich foods, drinks and

desserts such as pizza, soda or ice cream, numerous (but not all) psychoactive substances, both legal and illegal, such as cigarettes, alcohol, cocaine or heroin, pornography and irresponsible sexual behavior, gambling, certain types of entertainment, such as shows, movies or literature designed for entertainment with poor added value (where the main cost that is incurred is the opportunity cost of not, e.g., seeing a movie that is both entertaining and educational, or not doing anything productive), some types of social media, and certain games. This list is not exhaustive, people vary in which goods they overconsume, and by no means does it mean that by a good being on this list, it is wrong to consume it in any amount.

Through these goods, we can better investigate the topic of intertemporal choice through a multiple-process perspective and provide more arguments against centrally computed utility, as I believe that every reader of this thesis knows the feeling when urges pull in one direction, while reason and willpower push in the opposite direction.

There seems to be a positive correlation between affective activation and steepness of discounting, meaning that those goods which elicit more affective reaction are discounted more steeply. **Ubfal (2016)** cites a research article from **Bickel et al. (2011)**, which found that addicts have higher rates for their drug of choice than for money, and an article from **Tsukayama and Duckworth (2010)**, who report that those who feel tempted by certain goods discount them more steeply than those who do not feel tempted by the same goods.

Reflective thinking vs. gut-reaction

There seems to be a noticeable difference between choices that rely on gut reaction, which are done in experiments similar to those on animals, and choices which require deliberate reflective thinking. **Ainslie and Monterosso (2003)** note that during experiments in which people were given choices to delay or expedite actions such as having to listen to aversive noise, look at pictures of celebrities, or being given access to juice, and which had durations in the scale of minutes, found steep discounting, which was relatively similar across subjects, mirroring animal studies. In contrast, they note that across choices which require processes such as conscious planning or calculation, discount rates across individuals seem to vary considerably more. Furthermore, they note that

when using money, researchers often couldn't find any discounting in the time-scale of minutes – while in the experiments that were similar to animal experiments, "in each of these cases, the value of a reinforcer was diminished by half or more in an amount of time on the order of a minute. This represents a difference of six to seven orders of magnitude [!] when compared to discounting of large hypothetical monetary quantities **(Ainslie & Monterosso, 2003, p. 41)**." These findings favor the notion of multiple-process models, such as hot (affective) & cold (reflective) **(Heshmat, 2015b)**, or model-free (habitual) & model-based (goal-directed) **(Story et al., 2014)**, which model choices, even in the most simplified form, as results of interaction of various processes, which sometimes have varying preferences. Furthermore, **Urminsky and Zauberman (2016)** mention **Frederick's (2005)** research, which finds that those who are more likely to answer incorrect but intuitively correct answers to certain questions display steeper discounting. For example, the question "a bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?" has a usual impulsive answer of 10 cents, but the correct answer is 5 cents **(Frederick, 2005)**. Frederick argues that those less likely to reflect on the questions and answer incorrectly have lower levels of reflective thinking, and therefore reports negative correlation between reflective thinking and steeper discounting.

Possible psychological mechanisms of time-inconsistent preferences

If we want to go beyond the notion of centrally computed utility using a non-exponential discount function, we need to review some of the proposed mechanisms, which would explain dynamically inconsistent behavior even in individuals without explicit discount curves.

Intertemporal prisoner's dilemma

Consider the example of a dinner at a restaurant, when we are trying to watch our weight. How often do we actually tell ourselves that we would rather eat a dessert than have a healthy weight? More likely, we will try to find excuses, which would allow us to eat the dessert today, while maintaining the goal of a healthy weight in the long run. Of course, this possibility is alluring, because from the perspective of our current selves, the option of obtaining both the immediate reward and the long-term goal trumps having to choose either one. Much of our choices are like this – there are many short-term rewards available over time, and the opposing long-term goal isn't dependent on any one specific choice. Instead, often, each choice realistically has only a small effect on the long-term goal, and the attainment of the long-term goal depends on many choices done over time. This can lead us to reach the conclusion that we can choose the immediate reward in the present, and abstain in the future. Only once we are placed in front of the same choice, we reach the same conclusion – to indulge now, and abstain in the future. We can even feel betrayed by our past self, since unless we are naive, we know that had we abstained before, it would be easier to abstain again – that by indulging, we decreased our self-efficacy, i.e., the sense of competence to achieve a goal **(Heshmat, 2015a, p. 230)**, and that in future similar choices, we will be more likely to come up with different excuses to indulge. This pattern of repeated myopic preference reversal is perfectly consistent with hyperbolic discounting, however, some authors have provided an extending view using game theory, specifically repeated prisoner's dilemma **(Ainslie & Monterosso, 2003; Heshmat, 2015b)**. Despite its usual use case to analyze interaction of

multiple agents, in the case of individual intertemporal preferences, it is used to illustrate the interaction between successive states of an individual. Specifically, it is argued that choices by the present self will serve as precedents for choices of the future self, and choosing indulgence, i.e., the immediate reward, is seen as defection, which can undo the effect of many instances of cooperation, and sets a precedent for future defection **(Ainslie & Monterosso, 2003)**.

Connectedness to future self

Our level of connectedness to our future selves, defined as the extent to which beliefs, values, goals and other defining features of our future and current identity overlap **(Urminsky & Zauberman, 2016)** can play a role in the extent to which we devalue future outcomes. Of course, a person that is disconnected from their future self to the point where future costs are not a "my problem" but a "somebody else's problem" will have less of an incentive to sacrifice current consumption compared to a person who considers even distant future costs as their own. **Urminsky & Zauberman (2016)** note that future outcomes are discounted more steeply in periods where connectedness to the present self declines more rapidly, compared to periods in which connectedness changes less.

Opportunity cost consideration

Urminsky and Zauberman (2016) present the possibility that time-inconsistent choices can arise not only when future outcomes are devalued. Same choices can occur in a person who values delayed outcomes, but fails to consider the opportunity costs of current spending, thus simply not knowing that by choosing to consume in the present, they will not be able to reach the delayed goal. A person who fails to consider opportunity costs will repeatedly choose to spend money instead of investing them not because they prefer current consumption over early retirement, instead, they simply are not aware that their level of current consumption will prevent them from retiring early. It is possible that consideration of future outcomes is connected to people's discount rates, and also that they are separate, and both low discount rates along with consideration of

the future need to coexist for far-sighted options to be chosen **(Urminsky & Zauberman, 2016)**.

Resource slack theory

Similar to flawed opportunity cost consideration, resource slack theory centers around the perception of available resources **(Urminsky & Zauberman, 2016)**. According to this theory, hyperbolic discounting can be caused by a flawed prediction about the surplus of resources available in the future. If someone perceives that in the future, he will have more "slack", i.e., a surplus of available resources such as disposable income, free time, or energy, he will discount current resources more steeply. This mechanism would explain procrastination, i.e., postponing actions and decisions which are unpleasant, boring, or otherwise requiring sacrifice of current resources. Because why study now, if my available energy, motivation and willpower is low, when I can study tomorrow, when I will surely have more energy, and therefore motivation and willpower?

Construal level theory

Similar to opportunity cost consideration theory and resource slack theory, construal level theory circles around an imperfect representation of future outcomes, this time by imagining delayed outcomes more abstractly and with less details.

> This theory suggests that people may form abstract representations of distant, future events than near, future events (Trope and Liberman, 2010). From a distant perspective, choices are made based primarily on global concerns (why), whereas from proximal perspective, those priorities are weakened and even reversed as local concerns become more prominent (how). As people get psychologically closer to the situation, their choices are increasingly influenced by more specific concerns. Thus, distance impairs our ability to identify specific details of the choice. **(Heshmat, 2015a, p. 211)**

Thus, by not focusing on the relevant details of distant choices, such as their opportunity costs or difficulty to attain them, we fail to evaluate them objectively.

Hot & cold empathy gap

This theory, also known as projection bias, is mostly relevant to goods with high affective activation (tempting goods) and is another reason for self-control problems **(Heshmat, 2015a, pp. 210–211)**. Hot and cold empathy gap posits that when planning future consumption, while using our "cold" system of thinking, we fail to fully understand the effect that the choices will have on our emotions and urges at the point of encounter, when the "hot" system is activated **(Heshmat, 2015a, pp. 210-211)**. Because of this, we fail to take precautions against preference reversal, because from afar, they do not seem necessary. **O'Donoghue and Rabin (2000)** distinguish between sophistication and naivety, where a person who is naive is not aware that this preference reversal will occur, and therefore suffers from projection bias, while a sophisticated person is aware of the occurrence of preference reversal, and can therefore adjust their behavior in advance. **O'Donoghue and Rabin (2000)** also show that sophistication might not always be good for a person – a naive addict may relapse later than a sophisticated one, because in advance, he expects that he will abstain, while a sophisticated addict who knows that at some point, he will relapse, may under some circumstances expedite this process and relapse early, because he has less to lose. **Urminsky and Zauberman (2016)** call the ability to anticipate future emotions **affective forecasting**.

Objectivity of time

As was already mentioned regarding the salience of rewards, some distortions in discounting can stem from overestimating and underestimating the durations of time intervals. **Heshmat (2015a, pp. 212–213)** points to the paper from **Zauberman et al. (2009)**, which finds that impulsive individuals experience time differently, specifically, they overestimate the duration of time intervals, thus discount delayed rewards more steeply. However, some level of time distortion is pertinent to all humans - the closer a time interval is, the longer it seems – a year does not feel four times as long as three months (**Zauberman et al. 2009**).

People's perception of time is also susceptible to framing. The way in which choice is presented influences discounting, so that choices, which are essentially the same, can

have varying rates of discounting **(Loewenstein & Prelec, 1992).** For example, **Loewenstein (1988)** found that those who bought a good with expected delivery after a year were willing to pay an extra 54 dollars for immediate delivery, but demanded a compensation of 126 dollars for a delay of one year when the same good was supposed to arrive immediately, thus suggesting the formation of a reference point **(Loewenstein, 1988).** Also, interestingly, dividing time-periods which are the same length in total to a different amount of sub time-periods influences discounting, because a time period that is divided into several smaller time periods is perceived as longer than if it is not divided **(Li et al., 2022).** People's choices also vary by presenting the time as a defined point (e.g., 3 weeks from now) or an interval (2–3 weeks from now), and are also dependent on whether a future date is specified, or implied by the delay (e.g., "July 6th" or "a month from now", which is also on July 6th), with less discounting for options that are exact rather than "between X and Y" intervals, and less discounting for specified dates than when using time intervals to describe delays **(Li et al., 2022).** Furthermore, specifying information about the day when the reward will be received, such as events, makes people more likely to choose the larger, later reward, which, as **Li et al. (2022)** argue, suggests an effect of episodic future thinking on delay discounting.

Moreover, time perception is also associated with mood states and levels of dopamine. It was demonstrated that time seems to go by faster both when we are having fun and also after consumption stimulants, including caffeine **(Heshmat, 2015a, p. 213)**, which can help explain why people with ADHD, especially when bored, display more impulsive and novelty-seeking behavior, and counterintuitively, can improve their hyperactivity symptoms by consuming psychostimulants.

Furthermore, **Heshmat (2015a, p. 213)** also mentions a finding from **Sayette et al. (2005)**, that smokers who were feeling a strong urge to smoke experienced time as passing more slowly, and **Kim and Zauberman (2013)**, mentioned by **Urminsky and Zauberman (2016)** found time perception changes when people were exposed to sexual cues, which suggests a relationship between affective activation and time perception, specifically, that under the influence of "hot system" or visceral urges, time passes more slowly, which increases impatience, and therefore the discount rate.

This distortable inner clock further complicates discussions of rationality and underlying discount functions. It explains much of self-defeating, time-inconsistent behavior, but it adds another factor which must be considered and is hard to control for, i.e., the non-linear perception of delay, which is also influenced by both psychological (e.g., framing, mood) and physiological (e.g., dopamine level, substance withdrawal) factors.

Uncertainty

In the real world, introducing delay into a choice also introduces uncertainty. Violations of stationarity and even consistency, such as preferring larger reward after 101 days over smaller reward after 100 days and preferring smaller reward now over larger reward tomorrow, can be explained by a heuristic to prefer certain outcomes over uncertain ones **(Li et al., 2022)**. For an agent with such heuristic, the choices are far from equivalent. From a distance, both choices are uncertain, and the difference between them seems marginal. When the smaller reward is immediately available, however, it is a choice between smaller, certain reward, and larger, uncertain reward.

Multiple-process perspective

As was already indicated in previous chapters, preference reversals can be explained by the existence of multiple decision networks, which process information in different ways and at different times **(Kalenscher & Pennartz, 2008)**. Behavioral economists tend to propose a dual-system model of decision making, in which two, essentially separate systems process information, and compete between each other to reach a decision. Usually, they divide these systems into a hot, affective system and a cold, deliberate system **(Heshmat, 2015b; Ruhm, 2012)**. In this model, the affective system is near-sighted (myopic) and focuses only on immediate consequences, while the deliberate system is far-sighted and also considers long-term consequences. The affective system is thought to be faster, less conscious, less cognitively taxing, and is thought to rely on emotions, impulses and urges to motivate action. Deliberate system, on the other hand, is thought to be slower, more cognitively taxing, and motivating action through willpower and deliberation. In some cases, no claims about the neurophysiology of those systems are made. In other cases, behavioral economists utilize the triune brain theory **(e.g., Ruhm, 2012)**, which was proposed by the neuroscientist Paul MacLean in the 1960s **(Sapolsky, 2017, p. 27)**.

According to the triune brain theory, the human brain consists of three distinct parts, which have been gradually adding on top of each other during evolution, and are largely independent of each other. "Reptilian" brain, evolutionarily oldest, is responsible for unconscious, automatic processes, such as breathing and digestion. Surrounding it is the limbic system, which is attributed to mammals, and is supposed to process sensory inputs into feelings and emotions, which then influence the reptilian brain. The limbic system is what some regard as the "hot" or affective system **(Ruhm, 2012)**. Evolutionarily youngest is the neocortex, containing also the prefrontal cortex, which is presumed to be the center of cognition and planning – the so-called "cold" system or deliberate system **(Ruhm, 2012)**.

While the triune brain theory is to this day popular both among the psychological community and the general public, much of its premises have been disproven by neuroscientific research **(Steffen et al., 2022)**. Its claims about evolutionary development

of the brain have been challenged. Instead of simply adding new structures on existing, unchanging ones, it is presumed that even the evolutionarily older parts have evolved over time and adapted. Furthermore, while the original triune brain theory essentially states that these individual parts are largely independent, they have been shown to be interdependent and working closely with each other **(Sapolsky, 2017, p. 28)**. For example, the limbic system is not purely an emotional center, nor is the cortex a purely cognitive center **(Steffen et al., 2022)**. Frontal cortex is even regarded as an integral part of the limbic system **(Sapolsky, 2017, p. 33)**.

Research to determine the existence of multiple decision systems has not yet provided a unitary answer. Some studies claim having found evidence for the existence of a "hot" system which activates only when immediate rewards are available, while other provide countering evidence by showing that these parts also activate during some experiments with only delayed rewards, depending on the design of the experiment **(for comprehensive overview, see Kalenscher & Pennartz, 2008)**. While we can say with high confidence that the notion of independent, strictly emotional limbic brain competing with separate, strictly cognitive cortex is an oversimplification at best and can be misleading, it can put us on the right track of thinking. It was, for example, proven that "primary [e.g., food, water] and secondary [e.g., money] reinforcements invoke different psychological mechanisms, and recruit at least partially different neural networks (Bassareo and DiChiara, 1999; Parkinson et al., 1999; Grimm and See, 2000; Gottfried et al., 2002a, b; Estle et al., 2007)**(Kalenscher & Pennartz, 2008, p. 299)**." This could be relevant for deciding between saving money or buying food. Furthermore, the difference between imagining and actually experiencing delay lengths and rewards can substantially impact their valuation and discounting **(Kalenscher & Pennartz, 2008; Story et al., 2014)**. This suggests that whatever the exact neurophysiological mechanism, a choice between a primary reward available in the near future (e.g., a donut) will have a different neural evaluation than that of a secondary reward available in the distant future (e.g., investing for retirement). Furthermore, as will be shown in following chapters, prior experience will also influence the mechanism of neural evaluation – such as never having eaten a donut before, as opposed to having eaten one before.

Willpower

Another compelling argument in favor of the multiple-process model is also that it largely mirrors our subjective experience. There really does seem to be a part of us that is impatient, hedonic, and completely dismissive of any future costs and sacrifices. It wants to feel good, and it wants to feel good now. Much of our lives seem to consist of fighting its power, and much of those fights, we lose. What makes us resist some of our urges and impulses, and give in to others?

Some argue that this, precisely, is the role of willpower **(Heshmat, 2015a, p. 200)**. There also seems to be a reason, why sometimes we give in to the temptation of immediately gratifying choice, and why sometimes we resist. That is because willpower can be understood as a limited resource, which is depleted throughout its use in a similar fashion to energy, and there are several factors which influence it. **Heshmat (2015a, pp. 204)** calls it the strength model of willpower. He presents several factors which influence willpower. Willpower can be depleted by making choices which involve conflict, such as responding kindly to rude behavior, but also those which involve resisting immediate gratification **(Heshmat, 2015a, p. 201)**. **Baumaister (2002)** shows that self-control was repeatedly found to be the poorest among those who were already required to perform a prior act of self-control.

Other factors which seem to tax or undermine willpower resources are decision making, cognitive load (e.g., calculating, memorizing), stress, negative affect (e.g., anxiety, depression, trauma, unfortunate life events such as death in family, being fired, or divorce...), alcohol, low levels of blood sugar, sleep deprivation, close proximity to temptation (cue exposure), or licensing, i.e., rationalizing and giving oneself excuses to indulge due to special circumstances, such as celebrations, treats after hard work, or expecting to start a period of abstention later **(Heshmat, 2015a, pp. 200–204)**. Some people also seem to "give up" on their willpower efforts after violating their rules once, resulting in streaks of indulgence known as binges **(Heshmat, 2015a, pp. 207–208)**.

The strength model of willpower predicts that it is depleted throughout the day and that the depletion is transitive, i.e., depleting our willpower by doing one task will influence all other tasks which require willpower.

Working memory capacity seems to operate in a similar fashion to willpower, and can be influenced by similar factors (e.g., cognitive load, sleep deprivation). It is defined as "the ability to selectively attend and remember goal-relevant information **(Heshmat, 2015a, pp. 204–205)**." It can be influenced by similar factors as willpower. Its role in successful self-control lies in the ability to actively represent goals and goal-relevant information in working memory, and it's been shown to influence discounting **(Heshmat, 2015a, p. 205)**.

Heshmat (2015a, p. 201) then points out the trap that can be posed by poverty, which can prevent poor people from being able to take the actions necessary to improve their lives, because their willpower, working memory and attention resources are already depleted due to the numerous hard choices and sacrifices which they have to make based on their monetary constraints. All people who need to exert additional cognitive resources can face the same issue – e.g., recovering addicts who abstain despite experiencing cravings, people on a diet, people suffering from anxiety, depression or anhedonia, or people with impulsivity disorders such as ADHD.

Based on the provided evidence, I argue that individual psychological and physiological differences and environmental factors such as availability of resources or presence of affective cues influence the cognitive and affective states of individuals, which in turn influence the level of temporal discounting of such individuals, as well as determine their level of self-control, i.e., an ability to avoid dynamic preference reversals. Some goods, cues, physiological states or actions evoke higher affective activation, which at the very least contributes to impatience and temporal myopia, and helps explain both hyperbolic discounting, dynamic preference reversals, and good-specific discount rates. The stronger the affective state is, the more cognitive resources such as willpower and working memory capacity are needed to exercise self-control, i.e., abstain from choosing the affectively more attractive choice in favor of the cognitively more attractive, responsible (rational) choice. This process depletes cognitive resources, which explains why at different times, different choices are made. Irregardless of the neurophysiological mechanism, subjectively, this can feel like an ongoing battle between multiple parts of the self, which validates the framework of the hot & cold system from a psychological perspective.

Reinforcement learning perspective

An interesting approach in modeling intertemporal preferences came from the field of computer science, specifically, from the field of reinforcement learning. By reinforcement learning, we understand a computational approach to learning from interaction with an environment **(Sutton & Barto, 2018, p. 1)**.

> The idea that we learn by interacting with our environment is probably the first to occur to us when we think about the nature of learning. When an infant plays, waves its arms, or looks about, it has no explicit teacher, but it does have a direct sensorimotor connection to its environment. Exercising this connection produces a wealth of information about cause and effect, about the consequences of actions, and about what to do in order to achieve goals. Throughout our lives, such interactions are undoubtedly a major source of knowledge about our environment and ourselves. Whether we are learning to drive a car or to hold a conversation, we are acutely aware of how our environment responds to what we do, and we seek to influence what happens through our behavior. Learning from interaction is a foundational idea underlying nearly all theories of learning and intelligence **(Sutton & Barto, 2018, p. 1)**.

The reason I chose to include reinforcement learning in this thesis is that, similarly to economics, it is built on the idea of an agent maximizing some kind of reward through interaction with the environment that is often complex, uncertain, or even changing over time. The development of reinforcement learning was inspired by psychological learning theories **(Sutton & Barto, 2018, p. 341)**, and many of the challenges and situations faced by reinforcement learning agents resemble the challenges of natural learning systems. However, computational reinforcement learning primarily seeks to solve computational problems with efficient algorithms, rather than to replicate or explain how animals learn **(Sutton & Barto, 2018, p. 341)**. Nevertheless, interestingly, some algorithms developed independently from the field of neuroscience have been found to closely mirror real neural mechanisms, most notably the role of dopamine in reward prediction **(Sutton & Barto, 2018, pp. 380–395)**. Reinforcement learning therefore became an inseparable part of

computational psychiatry, a field that seeks to help understand mental disorders through mathematical and computational methods **(Sutton & Barto, 2018, p. 410)**.

Reinforcement learning models have already been used to model intertemporal preferences and to help explain phenomena associated with myopic, non-exponential reward discounting **(e.g., Kurth-Nelson & Redish, 2010; Fedus et al., 2019)**. A comprehensive analysis of contemporary research of intertemporal choice using reinforcement learning and computational psychiatry is outside the scope of this thesis.

In this thesis, I would like to specifically focus on a framework provided by **Story and colleagues (2014)**. They point out the resemblance of habitual and goal-directed learning to model-free and model-based reinforcement learning. Essentially, a model-based agent has some kind of a model of an environment in which he operates, and before he makes a decision, he searches through possible future states after each possible action. This model-based learning resembles goal-directed behavior of people and animals, and the agent's actions are highly sensitive to the change in the environment. A model-free agent, on the other hand, operates through experience. By aggregating past experience, he assigns value to each possible action, but has no explicit representation of the future state of the environment **(Story et al., 2014)**. There have been identified both parts of the brain involved in habitual learning, as well as goal-directed (model-based) learning in mammals **(Sutton & Barto, 2018, p. 407)**, and while it is not exactly known how these mechanisms arbitrage to reach a decision, it is assumed that behavior is initially model-based, and through experience becomes more habitual **(Story et al., 2014)**.

This mechanism makes sense, because model-free, habitual learning allows us to free our cognitive resources, as it is computationally less taxing. Instead of going through our model of the environment each time we need to take an action, we simply act out of habit, i.e., a simple stimulus-response pair learned by past experience. But, this comes at a cost. Firstly, deeply ingrained habits are much less flexible to changes in the environment. An experiment on rats showed that after overtraining them to press a lever (500 rewarded presses) to receive food, the rats continued to press a lever long after its value was devalued, something which didn't occur with rats who only did 100 rewarded presses **(Sutton & Barto, 2018, p. 368)**. A human example of this could be a driver repeatedly slamming the brake on a car with an automatic transmission with his left foot, because

that's where the clutch has always been, a behavior which we wouldn't expect from a never-before driver, for whom driving cannot yet be habitual.

There's another issue with this mechanism, much more relevant to the thesis. We assume that habits, i.e., specific stimulus-response pairs, become reinforced after being rewarded, and weakened after being punished. However, if this learning is model-free, with added delay, there emerges something which is analogous to the "credit assignment problem" in reinforcement learning – identifying the specific action to assign credit for success or blame for failure, if it could have been any of the numerous previous actions **(Sutton & Barto, 2018, p. 385)**.

Put simply – how quickly would a child learn to not touch a hot stove, if the pain occurred after a week? How could a dog associate getting a treat with patiently waiting before crossing a road, if it receives the treat only at the end of the day? Or more analogous to the hedonic overconsumption problem, if a certain behavior was reinforced with a "poisoned" treat, which would provide immediate reward, but caused discomfort the day after, could we expect the dog to associate the behavior or the treat with the delayed discomfort?

Story et al. (2014) use a simplified example of deciding between eating a cookie now, providing immediate value, or retaining weight, providing value after a delay. While for the model-based system, it could be preferable to abstain from eating the cookie in order to retain weight later on, the model-free system, provided that it has already experienced eating a cookie, will see no value in abstaining, because it cannot experience the benefits of such a decision. This framework seems consistent with empirical findings. In animals, it has been established that acquisition of new behaviors is less efficient with larger delays in feedback **(Schlingler & Blakely, 1994)**.

More importantly, it has been found that even in adult humans, pharmacokinetics of a substance, i.e., its activity in the body influences addiction. The faster a drug reaches the brain, the more of it reaches the brain, and the shorter the duration of its action, the faster it causes addiction, i.e., habituation of its use **(Allain et al., 2015)**.

To clarify, I do not mean to imply that physiologically, these processes are separate, or that by habituation, it is possible to "turn off" goal-directed, model-based processing. As **Sutton & Barto (2018, p. 409)** state, "[s]ummarizing the situation, Doll, Simon, and

Daw (2012) wrote that 'model-based influences appear ubiquitous more or less wherever the brain processes reward information,' and this is true even in the regions thought to be critical for model-free learning."

Nevertheless, although simplified, this mechanism explains why "tempting goods" are hard to resist in the present, while their consumption in the past is often an object of regret, and their future consumption is often underestimated. It could as well be a part of what behavioral economists call the hot & cold system. It also predicts myopic preference reversals without a need for universal non-exponential discount function **(Story et al., 2014).** It would also explain the extreme dichotomy between the short-term and long-term goals of an addict, as well as extremely myopic behavior which often leads to the loss of freedom, relationships, wealth and even life. Following citations contain explicit language, but it's vital to illustrate how it feels when the force of habit is too strong.

> [...] I end up at my bank and for the second time I game the ATM machine. The bank I use allows for deposits in ATM's but the deposits are in envelopes meaning they don't get checked until a couple days later and can only make $100 available today, fine with me. I grab an envelope, insert a blank check and tell the ATM it's for $100. I learned this trick the other day and like any junkie, I won't stop until this well is tapped out. In goes the blank check, out comes the $100, and down the drain goes my credit and bank account. Soon my account will be -$300 but who the fuck thinks about tomorrow? I sure don't. [...] **(Nate, 2011).**

Nate then goes on to drive while in the state of withdrawal and for the 100 dollars he just fraudulently took out from his bank account in his own name, he buys heroin and oxycodone, enough to last him one day.

> Sometimes it makes me happy. That's all I could think of, the only morsel of a reason I could cough up for why I continue time and time again to put myself through this shit. Why I severed all meaningful relationships in my life. Why I rip people off. Why I willingly peddle just another piece of my soul for one more glimmer of that brown womb. It's why I continue to push today into tomorrow, blending the promises of yesterday with the anxieties of the now. That's really what all this is about, avoiding

the now. Preventing the inevitability of direct experience from making its way into my frontal cortex. Just a crumb of the brown powder and the worldly noise dissolves along with it into that spoon. The rising vapors fill my nostrils triggering a primal response deep within my being. [...] **(Nate, 2020)**.

Economic, health and life implications

Intertemporal choices are a key component which forms the lives of individuals, and on a large scale, influences the wellbeing of nations. From the point of macroeconomics, intertemporal self-control issues of individuals lead to significant costs to society. Some costs can be estimated. For example, in 2016 alone, the medical costs associated with obesity in the U.S. were estimated to be $260.6 billion **(Cawley et al., 2021)**. **Heshmat (2015, p. 194)** then cites research from **McGinnis and Foege (1993)**, which estimates that 40% of deaths are attributed to poor self-control. However, to say that accurately estimating all of the global costs associated with self-control issues is difficult would be an understatement. We will instead focus on the individual implications of myopic preferences, especially to account for the individual differences in discounting and how those differences influence the wellbeing of individuals. I chose not to separate this chapter into multiple chapters, because often, the implications of myopic behavior are inseparable and interconnected. For example, a smoking habit not only costs money, but also increases the risk of future negative health outcomes **(Urminsky & Zauberman, 2016)**, as well as likely reversibly further increases monetary discounting **(Story et al., 2014)**. Smoking has even been used as a proxy variable of higher impatience in some studies which investigated the relationship of impatience to job search intensity, as well as to initial wages and subsequent wage growths **(Urminsky & Zauberman, 2016)**.

Health

In terms of health, steeper discount rates have been associated with less responsible behavior – "[f]or example, higher rates of temporal discounting have been shown to be related to a reduced likelihood to check blood pressure, obtain cholesterol testing, attend dental visits, exercise, receive flu shots, engage in safe sexual behavior, and be medically adherent (Bickel, 2015) **(Keidel et al., 2021, p. 2)**." **Keidel et al. (2021)** also mention higher rates of discounting being associated with other irresponsible behavior, such as not wearing a seat belt, or texting while driving, which can have serious health-related consequences. Probably the most significant link between temporal discounting and health-related behavior is that of substance use. Numerous studies found a consistent

relationship between monetary discount rates and smoking, alcohol consumption, or illicit substance use **(Story et al., 2014)**. Most notably, while higher monetary discount rates seem to predict a higher risk of addiction, being in the state of addiction likely further increases the steepness of discounting **(Story et al., 2014)**.

Saving and spending

Present-biased discounting predicts numerous behaviors which would be anomalous from the framework of exponential discounting. Firstly, it predicts that current spending will closely track income **(Laibson, 1997)**. Specifically, the more myopic a person is, the more will their current consumption track even predictable changes in income, instead of smoothing out consumption, such as saving for retirement, which was confirmed empirically – **Hurst (2004)** found that a subset of consumers who "followed near sighted consumption plans" had lower than predicted wealth during retirement. Moreover, those aware of their propensity to overconsume in the present will likely actively seek out illiquid investments as a form of precommitment against early withdrawal of these funds **(Strotz, 1955)**.

There is even evidence which suggests that people place a premium on the inability to withdraw funds prematurely – for example, a financial instrument that was once popular were so called Christmas saving clubs, to which people contributed weekly, received no interest on their deposits, and could only withdraw their money after a year, at the start of Christmas shopping season **(Thaler & Sunstein, 2008, p. 48)**. Another evidence that people are willing to pay a premium to control their future spending is the phenomenon of mental accounting. Although money is by definition fungible, i.e., substitutable, people often keep paying off loans with high interest rates while also having saved up money with low interest rates **(Thaler & Sunstein, 2008, pp. 49–52)**.

People with higher discount rates have also been found to be more likely to default on their loans and have lower credit scores **(Keidel et al., 2021)**, as well as to require more rapid return on investment into more energy-efficient equipment **(Urminsky & Zauberman, 2016)**.

Governments are also at least implicitly aware of people's limited ability to account for future consumption, which explains why they often incentivize retirement savings accounts by tax exemptions and even by direct contributions and disincentivize their early withdrawal before a certain age **(Laibson, 1997)**, as well as incentivize the adoption of energy-efficient equipment.

Employment and educational attainment

While the causal relationship remains unclear, steeper discounting and impatience has been associated both with lower professional and educational attainment. It's been found that people with higher discount rates are more likely to take a temporary job, which can be seen as trade-off between short-term monetary gain and future career advancement, and that more impatient individuals are more likely to have a lower initial wage, slower wage growth and are less likely to invest in their current job **(Urminsky & Zauberman, 2016)**. Likewise, higher rates of temporal discounting are correlated with poorer grades and a lower level of educational attainment **(Urminsky & Zauberman, 2016)**.

In conclusion, myopic preferences, such as impatience, propensity to overvalue the present, or to neglect future rewards and costs seem to be a crucial component which influences the wellbeing of individuals in all areas of life. It can therefore be in the best interest of individuals as well as organizations to know possible strategies, which could help combat myopic preference reversals. We will review some of these strategies in the next chapter.

Strategies for long-term utility maximization

The previous chapters made several points clear. Myopic preferences, dynamic preference reversals and struggles to exercise self-control are a phenomenon which relates to everyone in our society, at least to some extent. From the perspective of large organizations such as nations, and for some individuals, they represent costs of extreme magnitude. It is therefore in the best interest of both individuals and the organizations they are part of to help them make decisions which they will not regret in the future.

As we have established, making decisions which serve long-term goals at the expense of immediate pleasure taxes cognitive resources, such as willpower and working memory capacity. We can therefore divide all utility-maximization strategies into two categories. Either they increase the ability or motivation of cognitive resources to override impulses, or they decrease the difficulty of overriding those impulses. These strategies can then be employed either by the individual himself, or by others – such as by public policy.

Ainslie (2001) identified four mechanisms of combating the reversal of preference in favor of short-term reward, as it becomes more immediate. These are extraphysic commitment, manipulation of attention, preparation of emotion, and personal rules.

Extrapsychic commitment

Extrapsychic commitment or **precommitment** essentially means physically, quantitatively, or qualitatively changing the choice so that objectively, the new choice is different from the original one, preferably making the larger, later reward seem like a better choice, or the only available choice. This can be done by either preemptively removing the short-term reward, or to somehow change its rewarding power or costs associated with it.

Examples of individual-based extrapsychic commitment strategies include substance abusers voluntarily consuming medicines which remove the pleasure of consuming their drug of choice, such as opioid-dependent individuals taking Suboxone, which prevents them from being able to “get high”, as it contains naltrexone, a compound which binds strongly to opioid receptors, but provides no pleasure. Alcoholics then sometimes take disulfiram (Antabuse), a medicine which impairs their ability to process

ethanol, thus making the alcoholic experience horrible hangover symptoms only minutes after having a drink. Interestingly, naltrexone, while originally used for opioid abusers, has proven effective also for treating excessive alcohol consumption, because it also limits the pleasure felt from drinking **(Volpicelli et al., 1994)**.

Other examples of individual extrapsychic commitment include gamblers voluntarily banning themselves from participation in betting, lotteries or entering casinos. Those with unhealthy eating habits can precommit by ordering a tailored meal plan – a "box diet", or by keeping sweets and snacks at a greater physical and temporal distance by not keeping their supply at home, having to go to a nearby store each time they want one. Another precommitment strategy for a healthier diet can be buying small servings of those sweets, snacks or sugary drinks, so that after finishing their allocated portion, they are not tempted to finish the whole package/bottle **(O'Donoghue & Rabin, 2000)**. A more extreme case of precommitment to a healthier diet is gastric bypass surgery. Health behaviors are, of course, not limited to just diet. People can, for example, precommit to exercise by purchasing a membership card instead of pay-per-visit in gyms, although, it has been shown that this on average costs them more per visit, than if they paid for each visit **(Urminsky & Zauberman, 2016)**.

People can also voluntarily raise the cost of the myopic option by telling others about their resolution, so that self-control failure will also mean losing face in front of others **(Ainslie & Monterosso, 2003)**. Furthermore, people can pledge money that they will donate in case they violate their resolution **(e.g., Giné et al, 2010)**

In the case of personal finance and wealth, on an individual level, consumers can choose less liquid investments, and research suggests that steep discounters do just that **(Urminsky & Zauberman, 2016)**. They can set up automatic payments for insurance, savings and investments, so that instead of having to expel effort to make those payments, they would need to deliberately stop those payments.

On an organizational level, extrapsychic commitment can take on the form of, for example, laws, restrictions, fines, taxes and tax exemptions, incentives, grants, or punishments. For pathological behaviors such as addictions, nations raise the costs associated with various substances by making their possession punishable. This influences both their monetary price, and other costs associated with them – such as the possibility

of losing freedom or social status. Another approach to limit substance use is to impose taxes on them – it has been shown that despite their addictive properties, consumers do respond to changes in substance prices **(Vuchinich & Heather, 2003; Chaloupka et al., 2003)**. Besides the omnipresent minimum age requirements, it is also possible to limit their availability to certain times of day, or the places where they can be purchased, such as Finland's approach to alcohol sales **(Karlsson et al., 2020)**. Gambling and lotteries are also commonly regulated in a similar fashion. Similar approaches can be made for unhealthy foods, such as Britain's sugar tax **(Colborne, 2016)**.

Governments can also improve the health of their citizens through extrapsychic commitment by incentivizing and subsidizing health check-ups, providing affordable health care, giving monetary incentives for blood donations, or making people opt-out instead of opt-in to be an organ donor.

While in the case of addictions, health and saving, people seem to have plenty of room for their own precommitment, the nation can also greatly influence the long-term wellbeing of individuals by incentivizing education and supporting those facing poverty, considering that both lower education and higher poverty were associated with steeper discounting, and individuals may have limited ability to precommit to not be poor, or to attain education they can not afford.

Organizational extrapsychic commitments do not have to be government-wide, organizations such as companies and schools can also promote long-term benefits, for example, by limiting access to unhealthy food for students or employees, by incentivizing employees to be physically active (e.g., gym membership as a benefit to a job), or by contributing to employees' retirement savings.

In conclusion, extrapsychic commitments can decrease the difficulty to exercise self-control by numerous mechanisms, such as reducing the affective activation properties of myopic rewards through increasing their temporal and physical distance, by removing the option to choose the myopic reward altogether, or by making the long-term reward more attractive. On the other hand, other designs can serve as motivators to expel more cognitive resources to avoid the myopic choice, such as imposing taxes, fines or punishments associated with the myopic choice. In accordance with the prospect theory, losses of resources already gained loom larger than possible gains of additional resources,

even in extrapsychic commitment **(Heshmat, 2015a, p. 242)**. For example, we should expect better success of smoking cessation in people who wish to stop smoking if we give them 100 dollars at the start of their programme and let them keep the money if they succeed, than if we promised them 100 dollars at the end of the programme, should they succeed. The gym membership case, in which people are better off paying per each visit however, suggests that the cost should be associated with a distinct, well defined action.

Nevertheless, both the efficacy and ethicality of criminalization of severely addicted individuals is debatable, as the strength of their habit, i.e., the affective pull, poses a real question, whether they are capable of fully understanding the consequences of their actions, and even if they do, whether it is in their power to avoid them without external help.

Interestingly, precommitment behavior is not observed only in humans. **Ainslie (1974)** observed that, some pigeons, when they were given smaller, immediate food reinforcement for pecking a key and a larger, delayed food reinforcement for not pecking a key, when given a chance, removed their option of pecking the immediate reinforcement key, thereby forcing themselves to wait for the larger, delayed reward. **Kurth-Nelson and Redish (2010)** propose that, for at least some models of hyperbolic discounting agents, the ability and motivation to precommit stems from the very nature of the model, suggesting that when given the option, it is natural for hyperbolic discounters to display precommitment behavior, even without higher reasoning, if the precommitment choice is highly salient.

Manipulation of attention

As was shown in this thesis, the way in which choice is represented and understood can highly influence its outcome. For example, the construal level theory says that outcomes which are far away are represented more abstractly than when they are near. From afar, committing to going to the gym can be easy, especially when we focus on the why and not on the how. The less specific and the more abstract a goal is, the less likely it is to be attained **(Heshmat, 2015a, p. 227)**. "A highly abstract goal may have no obvious behavioral expression (e.g., to get healthy) **(Heshmat, 2015a, p. 227)**." There are several

factors which increase the chance of us following through with a long term goal. A known acronym for goal setting, SMART, means that goals should be specific, measurable, attainable, relevant, and time-bound. That way, we can shift our attention towards the how, divide our goal into specific subgoals, and have a better way to hold ourselves accountable.

Furthermore, having more than one goal can cause decision paralysis, because most times, attaining a goal requires time and cognitive resources, and it may not be feasible to achieve both, or at least not at the same time. Being able to focus on one goal at a time can therefore increase our chances of attaining it. Moreover, sometimes, it may be necessary to abandon a goal altogether.

> Researchers point out that striving for goals that cannot be brought to completion leads to a host of negative outcomes including negative affect, anxiety, and diminished well-being. In the same vein, those who disengage from unattainable goals experience heightened well-being. Unfulfilled and failed goals from which a person does not disengage create rumination—repeated and often intrusive thoughts about the incomplete goal. Often, they emerge in our dreams **(Heshmat, 2015a, p. 227).**

Over time, deliberate attention to a goal becomes more habitual, and repeated actions towards that goal become easier. Furthermore, not all attention is conscious, and goals can be triggered unconsciously by the environment.

> Thus, one way to avoid self-control failures of a particular sort is to avoid the company of those who suffer from such failures. Being around prompt, hardworking high achievers is one of the best ways of becoming prompt and hardworking, while associating with slackers is a good way of becoming a slacker **(Heshmat, 2015a, p. 228)**.

Furthermore, people make different choices based on what they focus on after encountering failure, including self-control failure. People with higher self-efficacy, i.e., confidence in their ability to attain a specific goal, are more likely to be task-diagnostic, i.e., ask themselves what they need to do now, while people with lower self-efficacy are more likely to be self-diagnostic, asking themselves what is wrong with them, which can

create negative affect, further undermining future self-control **(Heshmat, 2015a, pp. 230–232).**

Heshmat also mentions a concept of regulatory focus, which divides people into those who focus more on how to attain gains, and those who focus on how to avoid losses, as well as a similar concept of self-discrepancy theory, in which individuals are motivated by an ideal self (their own aspirations) and an ought self (their perceived obligations) **(2015a, pp. 234–235).** He argues that it is preferable to be focusing on attaining gains and becoming an ideal self over focusing on avoiding losses and becoming an "ought" self. Firstly, prospect theory tells us that people tend to overestimate the negative value of losses **(Kahneman & Tversky, 1979)**, so those especially dreading losses likely stray further from objective valuation needed for rational choice. Secondly, those focusing on avoiding losses and acting the way they think they are obligated to act are more likely to experience anxiety and negative affect, which taxes cognitive resources and further undermines self-control **(Heshmat, 2015a, p. 236).** As a whole, focusing on positive aspects of choices and acting on one's own aspirations makes one more likely to experience intrinsic motivation, which has been linked to creativity and vitality **(Legault, 2016).**

On an individual level, attention is sometimes undividable from emotion, which will be discussed in the following section. We can conclude that attention plays an important role in intertemporal choice both on conscious and unconscious level. Some manipulation of attention strategies are attainable through simple access to information, such as informing oneself on the rules for effective goal setting and self-monitoring, or by changing one's environment to modulate unconscious attention. However, some individual differences in attention can be linked to people's personalities, and shifting attention towards approaches which facilitate self-control can take considerable time and resources, such as therapy **(Heshmat, 2015a, p. 145–146).**

On an organizational level, behavioral economists are also familiar with the concept of "nudges", an approach which seeks to use people's deviations from normative economic theories to their advantage. For example, people tend to stick with the default option **("status quo bias"; Thaler & Sunstein, 2008, pp. 8)** – e.g., restaurants and cafeterias can increase the rates of people choosing water over soda with their meal, if

they make it the default option, the same can be said for healthier side dishes over fries. There are numerous possible designs and implementations of nudges. Essentially, they all work by making it easier to make the more desirable choice from a long-term perspective while honoring the individual freedom to choose otherwise, hence the name "nudge" and not a "shove". Some nudges could also be considered a form of extrapsychic commitment, because by going against them we may have to face additional non-monetary costs, such as having to spend time and energy to change our choice, or the emotional cost of going against what the majority is doing. Other nudges, which only for example change the order of foods on the menu, are only manipulating one's attention.

Preparation of emotion

It can be difficult to separate attention from emotion, since attention is also closely related to our emotions and affective state, and therefore discounting. For example, paying attention to affective cues can trigger cravings or desires, and as was shown, people are more impulsive in such states. Businesses of course know this, and readily use this to their advantage, often to the detriment of their consumers **(e.g., the social media "infinite scroll", Rixen et al., 2023)**. Governments, on the other hand, try to mitigate this in some cases, by banning or limiting the advertising of certain products, such as cigarettes and alcohol. Furthermore, governments often try to trigger opposite emotions, e.g., by putting frightening images on cigarette packs **(Heshmat, 2015a, 241)**.

Overall, we can divide the influence of emotions on intertemporal choices into two categories. First would be short-term emotions, urges and impulses, affective states which activate after the introduction of affective cues which create temptation. Internal and environmental cues, which make the object of desire being thought about in a "hot" way, such as smokers currently craving a cigarette **(Sayette et al., 2005)**, children being in front of candy **(Metcalfe & Mischel, 1999)**, alcoholics seeing other people consume alcohol **(Heshmat, 2015a, p. 206)**, can overshadow long-term goals and have been linked to both steeper discounting and self-control failure. Also, acute stress has been linked to the activation of hot-system processing, possibly to ensure survival in life-threatening situations by quick, instinctual action **(Metcalfe & Mischel, 1999)**.

In contrast, some emotional states can be extended and even chronic. **Heshmat (2015a, p. 208)** points out that negative affect is a significant trigger of self-control failure. Depression can influence intertemporal choice through numerous mechanisms. Depressed people may consider it futile to exercise self-control due to their feelings of worthlessness. They may also feel incompetent, undermining their self-efficacy. They may feel unable to exert the effort needed for self-control, and they may even find that even if they do, things such as exercise do not bring them joy. They tend to ruminate and fixate on their negative emotions, and can view certain myopic behaviors, such as substance use, as a form of escape **Heshmat (2015a, p. 208)**. Same goes for other types of emotional distress, such as anxiety. **Metcalfe and Mischel (1999)** point out that chronic stress can cause the "hot" system to be systematically more activated.

> In sum, emotional distress causes a behavioral shift toward immediate improvements in mood, and so people make poor decisions. Thus, the impulsive behavior may represent a strategic attempt to regulate affect (Tice et al., 2001). In other words, impulsive behavior may be viewed as a rational attempt to address a pressing concern **Heshmat (2015a, p. 209)**.

Both organizational and individual strategies that are designed to improve long-term utility maximization will be specific with regards to the duration of emotion or affect that they try to address. Examples of emotion preparation strategies of short term affect lapses in self control include setting "if-then" rules as contingency plans **(Heshmat, 2015a, pp. 256–258)**, in which a person sets a plan of action for when they feel an urge to do something, e.g., "if I catch myself craving a cigarette, I will do a breathing exercise". It can also be helpful for a quitting smoker to remind themselves that the cravings are usually very short-lived **(Heshmat, 2015a, p. 186)**. Therefore, part of emotion preparation for short-term urges stems from becoming sophisticated rather than naive **(O'Donoghue & Rabin, 2000)**, in the sense that we will be able to predict our hot-state emotions while in a cold state. **Ainslie (2001, p. 78)** also mentions the importance of putting emotional significance to the long-term goal. While being tempted by a dessert which would violate our diet, we can therefore instead focus on the emotional aspect of being proud of our

body, or on the guilt or shame that we will feel if we fail to attain it due to poor self-control.

Chronic emotional distress requires a different approach. Similar to poverty, it can often make myopic choices seem rational. Based on the specific affliction, individuals can seek specialized treatments, such as psychotherapy or pharmacotherapy.

Personal rules

Another strategy to combat myopic preference reversals is also one most relevant to willpower. It was found that both for humans and animals, bundling rewards together influences the choice that is made. For example, students who preferred small, early amounts of money when choosing individually for each pair of smaller-sooner/larger-later have shifted their preference towards preferring larger, later payoffs when choosing a series of five payoffs at weekly intervals, and the same has been found to hold for rats who were offered series of squirts of sucrose **(Ainslie, 2001, pp. 82–84)**. Personal rules utilize the same idea. While choices can be different every time, our previous choices largely influence our current choices, and our current choices largely influence our future choices. **Ainslie and Monterosso (2003)** provide a thought experiment. If a person that is trying to quit smoking was to know with certainty, that starting tomorrow, they will become smokers irregardless of their current choice, they would be better off smoking a cigarette today, rather than not smoking. If they were to know with certainty that starting tomorrow, they will become non-smokers, irregardless of their current choice, they would also be better off smoking a cigarette today. Only if their choice today was to act as a predictor of their future choice could they be better off abstaining from smoking today. **Ainslie and Monterosso (2003)** also mention the finding of **Kirby and Guastello (2001)**, who found that by merely suggesting to students that their current choices might serve as predictors of their future choices, their preferences shifted towards the larger, later reward, although not to the same extent as when choosing series in advance. A possible reason for that is while choosing a series of rewards is also a form of precommitment, choosing *as if* we were choosing a series of rewards requires willpower. The strategy of utility maximization using personal rules therefore works by choosing according to

long-term principles, instead of a case-by-case basis. A person that is tempted to violate this principle is therefore required to use his cognitive resources, especially willpower, to override his impulses and not give in to temptation.

This strategy can be extremely effective and in many cases, is essential to ensure self-control, because no strategy is universally applicable, nor unconditionally effective. However, two things must be mentioned. Firstly, "**rationalization, not proximity, is the most notorious threat to willpower (Ainslie & Monterosso, 2003, p. 45).**" As was discussed in the intertemporal prisoner's dilemma, in many everyday choices, from any point in time, it seems like an optimal strategy to indulge now, and abstain later. However, we are aware that current indulgence which goes against the rules that we set increases the chances of future indulgence, undermining our long-term goals. We are therefore motivated to find a way in which the current situation is somehow different, so that indulging does not violate our rules and isn't perceived as defection. People's ability to rationalize indulgence is not to be underestimated. We all know the classic "I'm stopping tomorrow." People can also rationalize indulging due to special circumstances, such as celebrations, rewards for hard work, being on holiday, or having the opportunity to try premium or exotic goods such as alcohol, which for some people can be perfectly acceptable, but for some can lead to addiction. It is not unusual for recovering addicts to relapse after years of sobriety **(Heshmat, 2015a, pp. 190–191)**. Even with the best intentions, rationalization can undermine our self-control efforts and can lead to a slippery slope, where after each time, it is easier and easier to violate our principles, until there is nothing left. It is therefore worth noting that rules which are clearly outlined are easier to follow than if there is no bright line between cooperation and defection **(Ainslie, 2001, p. 104)**. This can be especially problematic with people who use food to relieve stress, because the line between acceptable amount of food and excessive amount of food is blurred.

Another danger to long-term wellbeing is compulsions, i.e., rules which are too rigid and which serve not the longest-range interest, but a medium-range interest. For example, few of us have never consumed any alcohol, few of us have never eaten a dessert, and I assume that most people would not precommit to not drinking ever again, or not eating dessert ever again. Problem is that some longest-term goals are hard to

define (e.g., being healthy), which limits the possibility of creating rules, and therefore they must be rephrased somehow to make them more exact, more countable – only by defining them in such ways, we change their meaning. Because acquiring long-range goals may require the use of intuition and judgment calls, people who are especially worried of succumbing to short-range impulses, such as addicts or anorectics, abide by a much stricter set of rules **(Ainslie, 2001, p. 151)**. People are sometimes even seeking help to free them from their "punitive superego", because they feel imprisoned by their self-control rules **(Ainslie, 2001, p. 151)**. **Ainslie and Monterosso (2003)** note that sometimes, people even precommit to indulge, for example, some alcoholics plan their drinking days in advance and go off their dose of disulfiram, and resume taking it after their period of drinking.

In conclusion, there are numerous approaches which can help individuals avoid goal-incongruent myopic behavior. They work by identifying the influential factors of intertemporal choice, such as awareness, attention, emotion, environment and cognitive resources and try to influence them in such a way that makes choices more likely which are consistent with our long-term goals. They often utilize the very psychological factors which by themselves make choice seem myopic, such as people's willingness to precommit and even their perception of slack, such as the Save More Tomorrow retirement saving plan **(Thaler & Benartzi, 2004)**, in which people precommit to allocate increasingly higher proportions of their wages with concurrent wage increases. Furthermore, they utilize other factors which have been found to influence people's decision, such as their preference to stick with the default option, or preferring to choose only from a limited amount of choices to avoid decision paralysis. Some of these approaches can be utilized by the individual, many others are, however, in the hands of larger organizations and governments which those individuals are a part of. While much of these approaches are purely economic, such as limiting one's choice or raising its cost, there is no doubt that psychological factors play a huge role in such decisions and need to be understood by those who wish to incorporate long-term utility-maximizing strategies on an organizational level. Special considerations must be made with regards to the limited ability of cognitive resources to override myopic preferences after being depleted and to the factors which deplete those resources. While individuals can, to an extent, influence

such resources such as willpower or working memory capacity, or influence the factors which deplete them, the limitations of willpower alone to change one's behavior must be taken into consideration when designing choice architecture and public policy. What comes to mind first are people with mental health issues, poor households, the homeless, and those with an addictive substance habit. In these cases, there is often high comorbidity, meaning that these behaviors occur concurrently. Furthermore, they seem to potentiate each other on both psychological and physiological level, often leaving those afflicted in a myopic trap with grim prospect of recovery, irregardless of the amount of criminalization, ostracization or punishment. I argue that it is therefore in the interest of many countries to rethink the strategies of dealing with such problems.

It is also worth pointing out that some people face the opposite issue of not allowing themselves to allocate a certain amount of resources to current consumption either because they compulsively save **(misers; Thaler & Sunstein, 2008, p. 52)**, or because they have developed a compulsion, a mid-range habit designed to fight the short-range habit instead of relying on pure willpower, in fear of giving in to temptation **(Ainslie, 2001, p. 151)**. Both authors argue that it may be in the person's longest term interest to allocate a certain amount of resources to current consumption, either in the form of a "fun account" that is meant to be spent on things which people enjoy, **(Thaler & Sunstein, 2008, p. 52)**, or, perhaps more controversially, than even an alcoholic could be better off if he incorporates some periods of drinking into his life **(Ainslie, 2001, p. 156)**. People have been found to increase their hedonic consumption when focusing on the long-term regrets of hedonic under-consumption **(Urminsky & Zauberman, 2016)**. It is therefore not my intention to even imply that one should feel shame or guilt for occasional indulgence, on the contrary, I believe that it is one of the things which make life worth living. A problem arises when consumption turns into overconsumption and when short-term interests are in direct conflict with long-term interests, endangering our development and wellbeing. This is when these strategies come into play.

Discussion and possible directions

Based on the provided evidence, it can hardly be refuted that the level to which an individual devalues future utility can have severe implications on the wellbeing of the individual as well as those around him. However, after reviewing theories of self-control and intertemporal choice from the perspectives of other fields, I am pessimistic in the ability of purely economic research to fully describe this phenomenon.

As was mentioned previously, steeper discounting has been associated with various irresponsible behaviors, such as fewer medical check-ups, irresponsible sexual behavior, or not wearing a seatbelt. However, consider, for example, an individual with a considerable distaste for uncertainty and risk, using a heuristic to avoid risk when possible. Such individual should, especially when assuming linear utility function of money, as is often the case **(Andersen et al., 2008)**, display high monetary discount rates. However, we would expect this risk-averse individual to also wear a seatbelt, as well as to attend dental visits or to listen to the instructions of his doctor. In contrast, an individual who discounts money steeply due to undervaluing his future wellbeing will also likely be the one not wearing a seatbelt or a condom. This suggests that different mechanisms are at play for myopic behavior in different individuals, and could explain poor correlations of discount rates, which have been found among different outcomes within individuals **(Story et al., 2014)**.

It is therefore likely insufficient to be talking about implications of steep temporal discounting in the limited sense of elicited discount rates and shapes of discount functions, and we should be more concerned with identifying the specific factors and processes behind one's neglect for the future. This suggests that should temporal discounting be used as a diagnostic tool, it will likely need to have not a form of a set of binary choices between monetary amounts, it will more likely resemble a psychological questionnaire, which would be used to unravel person's perceptions about the future (such as perceived opportunity costs of current behavior, anticipated resource slack in the future, salience of future rewards and costs, interconnectedness with future self), his level of affective forecasting (sophistication about his future emotional states, projection bias), his level of impulsivity (especially lack of premeditation), as well as to identify the

presence of risk factors which increase the chance of using maladaptive behavior as a coping mechanism, such as mental and other health issues which undermine willpower and other cognitive resources needed to exercise self-control and patience or even make the use of hedonic goods seem rational to the individual (e.g., anxiety, impulsivity disorders, learning disabilities, anhedonia, depression, PTSD, chronic pain). Of course, it should identify substances that the person already uses and his level of dependence and habituation based on factors such as route of administration, frequency of use and amount of use. It should also identify person's socioeconomic circumstances, such as his education, wealth, debt, credit score, wage, financial literacy, health insurance, and expected changes in future income (such as how close he is to retiring or what are his career prospects), as well as the person's environmental factors – the nature of his social circle, friends, family, the amount of available support from both emotional and financial perspective, level of safety of the environment as well as available opportunities for healthy behavioral expression, such as gyms, parks or libraries. I argue that only after such comprehensive analysis of one's current wellbeing as well as prospects of future wellbeing can any conclusions be made about the discounter's "phenotype" and tailored solutions proposed to increase the future wellbeing of an individual, for example by financial literacy lessons, financial precommitment devices, pharmacotherapy or substitution therapy, psychotherapy, help with expanding one's education or professional skills, or help with changing one's environment.

This thesis can therefore be a starting point leading into several directions. First would be to investigate psychological and economic diagnostic tools which would provide a comprehensive depiction of one's current and future prospects. If not already developed, a framework or a battery of tests could be proposed, which could be used to provide a complex insight into one's current and future wellbeing prospects by utilizing insights both from economics and psychology. If there exists already a test battery which would provide an extensive view into one's determining factors of future utility factors, which would allow to identify the "phenotype" of an individual's propensity to discount future, further research can be made into its efficacy, as well as possible integration into the health care system of a nation or into public policy. From an economic perspective, economic simulations can be made which would compare state's expenditures if the states

utilize some of the possible strategies, such as providing subsidized housing to the homeless (e.g., Finland's Housing first policy; **Juhila et al., 2022**), compared to the costs incurred by not adopting these strategies, such as increased medical costs, increased costs to fight criminality, as well as the costs of incarceration of some of these individuals.

Another possible direction could be investigating the economic interaction between the consumers and sellers of hedonic goods. As **Ubfal (2016)** pointed out, in a perfect market, the discount curves of goods should mirror the discount curves of money, because additional goods can be sold for money, and if their level is insufficient, they can be bought for money. Furthermore, he found that individuals who often participated in trade, in his case mostly farmers who sold their surplus to make a living, discounted various goods on average closer to their discount rates of money, as well as that when reminded of trade opportunities, people in general discounted various goods closer to the way they discounted money. In the case of hedonic goods, however, the myopic habituation model predicts that consumers of hedonic goods, especially if they are naive (have poor affective forecasting) will fail to fully predict their future consumption. And even if they do, few of them will precommit to their future consumption, because from afar, they would prefer abstaining. This makes the interaction uneven, in the sense that a financially literate seller of a hedonic good will discount the hedonic good in a similar way that he discounts money, i.e., without present bias, while a consumer of a hedonic good will discount such good oftentimes with a significant present bias. In such an interaction, the seller and the consumer have conflicting interests. It is in the consumer's best interest to have as little present bias as possible to avoid self-control issues and dynamically inconsistent behavior. In contrast, it is in the seller's best interest to increase the consumer's present bias. It is therefore no surprise that salespeople are often taught to sell by appealing to people's emotion, rather than through facts about the good. In the real world, there are many examples of using psychological factors influencing present bias and addictivity of a good to the benefit of the seller **(e.g., advertising addictive substances – alcohol: Saffer, 1991; cigarettes: Moran et al., 2021; the social media "infinite scroll", Rixen et al., 2023)**. A competition between sellers could suggest that this "arbitrage opportunity" will be lessened as more sellers enter the market, on the other hand, it could mean that other sellers are forced to also ensure the elicitation of

present bias on the consumer in order to stay competitive. This suggests that with certain goods, there is a significant incentive for the sellers to act against the long-term interest of their buyers, suggesting a formation of an "internality" - a within-person externality **(Heshmat, 2015a, p. 270)**. It is therefore in the interest of nations to identify goods which cause this form of externality, because without accounting for it, sellers of hedonic goods are incentivized to exploit it to not go out of business.

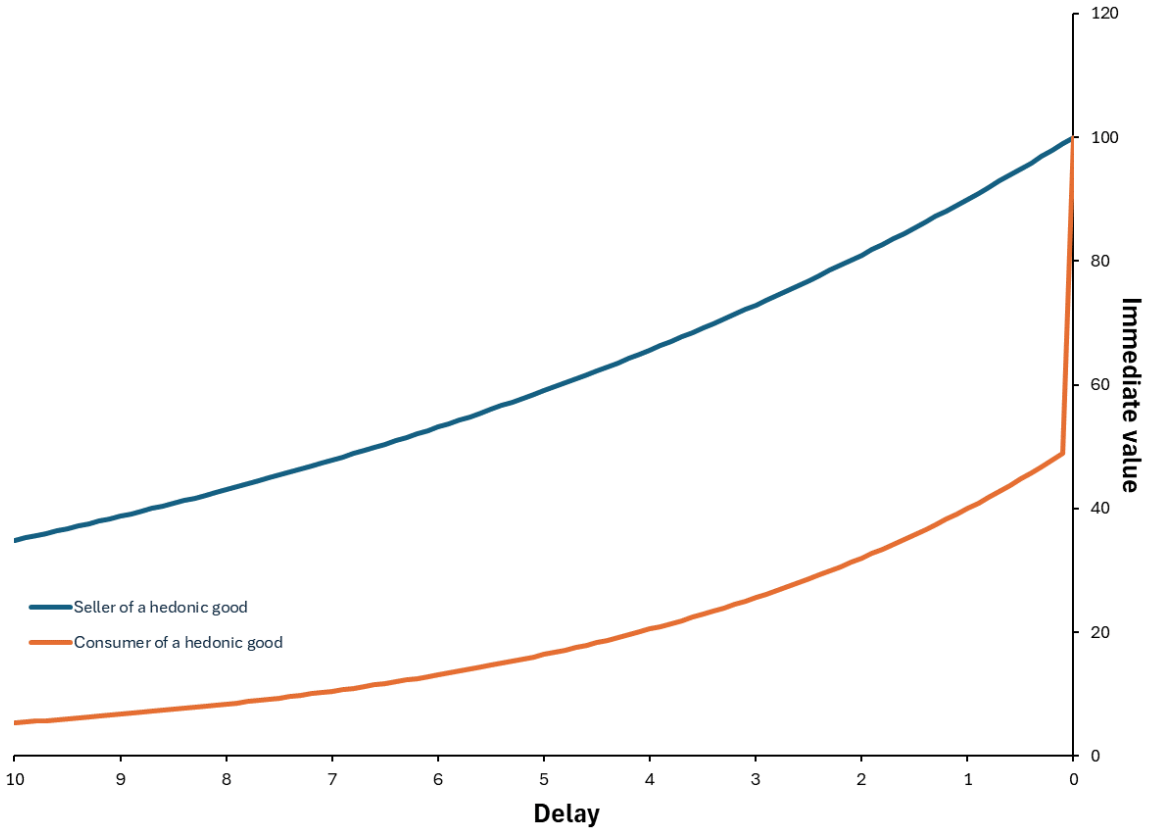

Figure 4.1: An illustration of a valuation of a hypothetical hedonic good in time from the point of view of the seller and from the point of view of the consumer.

It could also therefore be meaningful to explore the ethicality of various forms of marketing of goods which can be viewed as hedonic, as well as whether the sellers should be forced by law to offer precommitment opportunities to their consumers – such as those who wish to use social networks to have the ability to disable some of their functionalities, or shoppers having the opportunity to buy groceries without having to walk through the alcohol isle, or see an advertisement for tobacco products.

Another direction could be to explore the marketing of the opposite type of goods, "investment goods," such as gym memberships, healthy meal plans, therapy, or investment instruments. Myopic preferences predict that people will under-consume them, so it could be beneficial to explore the ways by which they can be effectively advertised, as well as ways which would increase customer retention.

The thesis also provides those in the addiction field with a perspective of both the role of deliberate choice as well as its influential factors in addiction. It can help them with designing intervention programs in the sense that it identifies numerous factors which influence choice, as well as shows the importance of what behavioral economists call the choice architecture. It can help them explain to their patients the role which deliberate choice and its influential factors have in addiction recovery. It shows that while a strong habit may be impossible to break by the sheer force of will, there are sets of conscious, deliberate, attainable actions that an individual can take one step at a time, which can greatly aid in his recovery. For example, an individual can be reminded that his habit will likely be less severe, if he chooses routes of administration which provide the same substance, but are less habit-forming – such as using nicotine patches over cigarettes. Of course, a smoker does not only have a nicotine habit, he also has a smoking habit, a puffing habit, a reaching for his cigarette pack when drinking coffee habit. An intravenous heroin user will also have a needle habit. Such is the nature of Pavlovian learning. And despite having his drug of choice, breaking such habits will require willpower. But this is precisely the role of deliberate choice in addiction, breaking an insurmountable obstacle into several smaller obstacles which can be overcome. The individual can then be informed about the effects which psychotherapy may have on his emotional state as well as on his ability to deal with setbacks in a healthy way. They can be reminded of how their environment and attention influence their choices, as well as how to prepare themselves for different emotional states. This approach avoids telling the patient that he is powerless and dependent on external help, which could also serve as an excuse for indulgence, as well as telling him that continuing his habit is his free choice, which is an oversimplification that can lead to undeserved guilt and shame.

Conclusion

When I picked the topic of the thesis, I had the illusion that somehow, I will be able to paint the full picture. What is now clear to me is that even after months of studying this topic, I am almost as far from fully understanding the reviewed phenomena as I was when I started. Not only that, I do not feel any more immune to the call of the Sirens, a metaphor depicting the irresistible pull of instant comfort, which comes at a great cost.

The main feature and added value of this thesis is its interdisciplinary approach. It shows that the topic can be approached using knowledge from several fields. Each of these fields can extend the collective understanding of the topic. Furthermore, researchers from other disciplines can often provide innovative and out-of-the box insights and solutions, simply because they used different patterns of thinking specific to their original field. A complete interdisciplinary integration of the topic may not yet be possible, especially due to the extremely complex nature of the brain's neurophysiology, which is yet far from being fully understood. However, I firmly believe that specialists researching the topic of intertemporal choice from any specific discipline, be it traditional economics and mathematics, behavioral economics, psychology, addictology, sociology, public policy, and even neuroscience and reinforcement learning, can benefit themselves by having some understanding about the topic from different disciplines, as well as provide new, innovative insights.

The interdisciplinary approach of this thesis is also its biggest weakness. By reviewing the topic from the perspective of several disciplines, in each of those disciplines, I was able to provide only an introductory overview. It's possible that some terms and concepts were insufficiently explained, which would require the reader to refer to additional sources to be able to finish reading the thesis, and despite my best efforts, I myself could have understood some concepts incompletely. It is therefore my responsibility to inform the reader that by no means does this thesis capture the phenomena of intertemporal choice and myopic preferences in their full complexities. To those who wish to gain a deeper understanding of the topic, I encourage you to review some of the sources that I cited, as well as refer to other relevant sources, which I inevitably missed.

Despite its limitations, I do believe that the goals of this thesis which were stated in the introduction were achieved, and that the thesis, as it stands, is of value to society. If nothing else, I would like both myself and the reader to remember one thing. Do not confuse comfort with happiness.

Summary

This bachelor thesis reviews the phenomena associated with intertemporal choice, such as myopic time-inconsistent preferences, goal-incongruent behavior and self-control failures using an interdisciplinary perspective. At first, it reviews the very basic concepts of utility and rationality, providing insights from the field of neuroscience and behavioral economics. It then reviews the original normative economic theory of intertemporal choice, exponentially discounted utility proposed by **Samuelson (1937)**, as well as time-consistency axioms which have been derived from it. This theory is then compared with findings from behavioral research on humans and animals, and it is shown in the thesis that exponentially discounted utility theory does not provide a true picture of actual economic behavior observed in humans. Alternative economic theories of temporal discounting are reviewed, specifically **Mazur's (1987)** hyperbolic discounting formula and **Laibson's (1997)** quasi-hyperbolic discounting formula. The view that intertemporal choice relies on a specific discount function is challenged due to their poor correlations and predictive power. Some limitations and challenges of traditional economic intertemporal choice research are then discussed.

The thesis then focuses further on empirical findings regarding intertemporal choice. It identifies factors which have been shown to influence intertemporal choice and steepness of discounting across individuals, and across different goods and outcomes within individuals. Seeking explanation for phenomena poorly predicted even by non-exponential discount functions, alternative mechanisms from a unitary discount function are reviewed through the lens of psychological research and reinforcement learning, with primary focus on multiple-process models. It was found that qualitatively different goods, as well as their temporal distance do indeed employ at least partially different neural mechanisms, and that there are numerous psychological and physiological factors which influence both steepness of discounting and the ability to exercise self-control. Most importantly, it was shown that cognitive resources such as willpower and working memory capacity influence the ability to override myopic impulses, and that using these resources depletes them and undermines future self-control efforts.

Using reinforcement learning, especially a model proposed by **Story et al. (2014)**, the thesis illustrates a possible mechanism to explain habituation of hedonic goods, i.e., goods which provide immediate pleasure and delayed costs, such as addictive substances. Specifically, it is argued that a computationally simpler learning system makes repeated goal-directed behaviors into habits to free up cognitive resources for new learning, however, the nature of hedonic goods makes it so that their habituation is reinforced by their immediate rewarding power, but is not then weakened by their delayed negative outcomes, because the computationally simpler learning system cannot associate those outcomes with the previous consumption. This creates discrepancy between habitual and goal-directed choice evaluation, which further taxes cognitive resources required to override the power of habit. In extreme cases, a situation arises in which it can be impossible for an individual to break his habit, which can have devastating consequences on his life and the lives of others. Furthermore, also relating to addiction, it was found that drug dependence could have both state- and trait-based influences, so that steeper monetary discounting increases the risk of becoming addicted, and the state of dependence then further increases monetary discounting.

The thesis then reviews some of the strategies to influence people's intertemporal choice towards long-term wellbeing and better self-control. It was found that through precommitment, manipulation of attention, preparation of emotion and personal rules, various strategies can be employed by individuals, professionals, organizations and governments to help people achieve better economic, psychological and health outcomes. Special consideration should be made when creating these strategies to account for the limitations of cognitive resources, making the right choice should be made easy, utilizing the very heuristics which have been identified, such as people's willingness to precommit or their preference to stick with the default option.

The thesis then discusses various possibilities for future research. Firstly, it could be beneficial to assess the discounting rates elicitation methods as well as psychological tests and questionnaires which are directly or indirectly relevant to one's current and future wellbeing, to assess whether there already exists a comprehensive framework which could be used to identify the individual's vulnerability factors as well as suggest tailored solutions for long-term utility maximization. After identifying or developing such

a framework, its possible modalities of adoption by governments and health systems can be investigated. It will also be of interest to investigate the economic implications of adopting such a framework and different strategies that it would imply.

Secondly, of interest to researchers could be the interaction between the sellers and consumers of hedonic goods. While it may be in the long-term interest of consumers to not be present-biased, it can be in the best interest of sellers to increase the present bias of consumers, they may be even required to do so if their competition is doing it as well. It is therefore likely that sellers of hedonic goods use attention, emotion and lack of precommitment to their economic benefit, while creating a negative externality to the consumer. It is in the power of public policy makers to mitigate this negative externality. In contrast, marketing of "investment goods", i.e., goods which incur immediate cost but provide larger benefit later, could also be of interest in future research.

Lastly, the thesis discusses the role of deliberate choice and the factors which influence it in addiction treatment and the extent to which those afflicted can use them both concurrently and independently of professional help.

References

Ainslie, G. (1974) Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior, 21*, 485–489.

Ainslie, G. (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person.* Cambridge: Cambridge University Press.

Ainslie, G. (2001). *Breakdown of Will*. Cambridge: Cambridge University Press.

Ainslie, G., & Monterosso, J. (2003). Hyperbolic Discounting as a Factor in Addiction: A Critical Analysis. In R. E. Vuchinich & N. Heather (Eds.), *Choice, behavioural economics and addiction* (pp. 35–69). Pergamon. https://doi.org/10.1016/B978-008044056-9/50043-9

Allain, F., Minogianis, E.-A., Roberts, D. C. S., & Samaha, A.-N. (2015). How fast and how often: The pharmacokinetics of drug use are decisive in addiction. *Neuroscience & Biobehavioral Reviews, 56,* 166-179. https://doi.org/10.1016/j.neubiorev.2015.06.012

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting Risk and Time Preferences. *Econometrica, 76*(3), 583-618. https://doi.org/10.1111/j.1468-0262.2008.00848.x

Banerjee, A., & Mullainathan, S. (2010). The Shape of Temptation: Implications for the Economic Lives of the Poor (NBER Working Paper No. 15973). National Bureau of Economic Research, Inc.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics, 7*(1), 22–53. http://www.jstor.org/stable/43189512

Baumeister, R. F. (2002). Yielding to Temptation: Self-Control Failure, Impulsive Purchasing, and Consumer Behavior. *Journal of Consumer Research, 28*(4), 670–676. https://doi.org/10.1086/338209

Bickel, W., Landes, R., Christensen, D., Jackson, L., Jones, B., Kurth-Nelson, Z., & Redish, A. (2011). Single- and cross-commodity discounting among cocaine addicts: The commodity and its temporal location determine discounting rate. *Psychopharmacology,* 1–11.

Callender, C. (2021). The Normative Standard for Future Discounting. *Australasian Philosophical Review, 5*(3), 227-253. https://doi.org/10.1080/24740500.2021.2112087

Catania, A.C. (1963). Concurrent performances: A baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior, 6,* 299-300.

Cawley, J., Biener, A., Meyerhoefer, C., Ding, Y., Zvenyach, T., Smolarz, B. G., & Ramasamy, A. (2021). Direct medical costs of obesity in the United States and the most populous states. *Journal of managed care & specialty pharmacy, 27*(3), 354–366. https://doi.org/10.18553/jmcp.2021.20410

Chaloupka, F. J., Emery, S., & Liang, L. (2003). Evolving Models of Addictive Behavior: From Neoclassical to Behavioral Economics. In R. E. Vuchinich & N. Heather (Eds.), *Choice, Behavioural Economics and Addiction* (pp. 71-94). Pergamon. https://doi.org/10.1016/B978-008044056-9/50044-0

Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 10,* 67-74.

Colborne M. (2016). Britain's "sugar tax" tackles obesity. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne, 188*(8), E134. https://doi.org/10.1503/cmaj.109-5260

Fedus, W., Gelada, C., Bengio, Y., Bellemare, M.G., & Larochelle, H. (2019). Hyperbolic Discounting and Learning over Multiple Horizons. ArXiv, abs/1902.06865.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*(4), 25-42. https://doi.org/10.1257/089533005775196732

Gaillard, A., Fehring, D. J., & Rossell, S. L. (2021). A systematic review and meta-analysis of behavioural sex differences in executive control. *European Journal of Neuroscience, 53*(2), 519-542. https://doi.org/10.1111/ejn.14946

Giné, X., Karlan, D., & Zinman, J. (2010). Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. *American Economic Journal: Applied Economics, 2*(4), 213-235. https://doi.org/10.1257/app.2.4.213

Glimcher, P. W., Dorris, M. C., & Bayer, H. M. (2005). Physiological utility theory and the neuroeconomics of choice. *Games and economic behavior, 52*(2), 213–256. https://doi.org/10.1016/j.geb.2004.06.011

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the experimental analysis of behavior, 4*(3), 267–272. https://doi.org/10.1901/jeab.1961.4-267

Herrnstein, R.J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243-266

Heshmat, S. (2015a). *Addiction: A behavioral economic perspective.* Routledge/Taylor & Francis Group.

Heshmat S. (2015b). Behavioral Economics of Self-Control Failure. *The Yale journal of biology and medicine, 88*(3), 333–337.

Hornby, A. S., Lea, D., & Bradbery, J. (2020). *Oxford advanced learner's dictionary of current English* (Tenth edition). Oxford University Press.

Hurst, E., & Lusardi, A. (2004). Liquidity Constraints, Household Wealth, and Entrepreneurship. *Journal of Political Economy, 112*(2), 319–347. https://doi.org/10.1086/381478

Juhila, K., Kirsi, S., Raitakari, S., & Ranta, J. (2022). Housing First: Combatting Long-Term Homelessness in Finland. In C. de la Porte, & others (Eds.), *Successful Public Policy in the Nordic Countries: Cases, Lessons, Challenges* (pp. 1-16). Oxford. Oxford Academic. https://doi.org/10.1093/oso/9780192856296.003.0024

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

Kalenscher, T., & Pennartz, C. M. A. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Progress in Neurobiology, 84*(3), 284-315. https://doi.org/10.1016/j.pneurobio.2007.11.004

Karlsson, T., Mäkelä, P., Tigerstedt, C., & Keskimäki, I. (2020). The Road to the Alcohol Act 2018 in Finland: A conflict between public health objectives and neoliberal goals. *Health Policy, 124*(1), 1-6. https://doi.org/10.1016/j.healthpol.2019.10.009

Keidel, K., Rramani, Q., Weber, B., Murawski, C., & Ettinger, U. (2021). Individual Differences in Intertemporal Choice. *Frontiers in Psychology, 12,* 643670. https://doi.org/10.3389/fpsyg.2021.643670

Kim, B. K., & Zauberman, G. (2013). Can Victoria's Secret change the future? A subjective time perception account of sexual-cue effects on impatience. *Journal of Experimental Psychology: General, 142*(2), 328.

Kirby, K. N., & Guastello, B. (2001). Making choices in anticipation of similar future choices can increase self-control. *Journal of Experimental Psychology: Applied, 7*(2), 154-164

Kurth-Nelson, Z. and Redish, A. D. (2010). A reinforcement learning model of precommitment in decision making. *Frontiers in behavioral neuroscience, 4*, 184. https://doi.org/10.3389/fnbeh.2010.00184

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics, 112*(2), 443–477. http://www.jstor.org/stable/2951242

Legault, L. (2016). Intrinsic and Extrinsic Motivation. Encyclopedia of Personality and Individual Differences. https://doi.org/10.1007/978-3-319-28099-8_1139-1

Li, S., Zhou, M., & Attia, A. (2022). A review of intertemporal decision making in neuroscience and psychology: Time perception, attentional resources, and emotion. In *International Conference on Management Science and Engineering Management* (pp. 652-666). Cham: Springer International Publishing.

Loewenstein, G. F. (1988). Frames of Mind in Intertemporal Choice. *Management Science, 34*(2), 200–214. http://www.jstor.org/stable/2632061

Loewenstein, G., & Prelec, D. (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. *The Quarterly Journal of Economics, 107*(2), 573–597. https://doi.org/10.2307/2118482

Mahalingam, V., Stillwell, D., Kosinski, M., Rust, J., & Kogan, A. (2014). Who Can Wait for the Future? A Personality Perspective. *Social Psychological and Personality Science, 5*(5), 573-583. https://doi.org/10.1177/1948550613515007

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In: M. Commons, J. Mazur, J. A. Nevin, & H. Rachlin (Eds), *Quantitative analysis of*

behavior: The effect of delay and of intervening events on reinforcement value (Vol. 5). Hillsdale, NJ: Lawrence Erlbaum Associates.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*(1), 3–19. https://doi.org/10.1037/0033-295X.106.1.3

Moran, M. B., Heley, K., Czaplicki, L., Weiger, C., Strong, D., & Pierce, J. (2021). Tobacco Advertising Features That May Contribute to Product Appeal Among US Adolescents and Young Adults. *Nicotine & tobacco research: official journal of the Society for Research on Nicotine and Tobacco, 23*(8), 1373–1381. https://doi.org/10.1093/ntr/ntaa275

Nate. (2011). All Roads Lead to the Same Dead End: An Experience with Heroin & Oxycodone (exp82799). Erowid.org. https:/erowid.org/exp/82799

Nate. (2020). Sometimes It Makes Me Happy: An Experience with Heroin & Naltrexone (exp111682). Erowid.org. https:/erowid.org/exp/111682

O'Donoghue, T., & Rabin, M. (2000). The Economics of Immediate Gratification. *Journal of Behavioral Decision Making, 13*: 233-250

Parthasarathi, T., McConnell, M. H., Leury, J., & Kable, J. W. (2017). The Vivid Present: Visualization Abilities Are Associated with Steep Discounting of Future Rewards. *Frontiers in Psychology, 8*, 289. https://doi.org/10.3389/fpsyg.2017.00289

Petry N. M. (2003). Discounting of money, health, and freedom in substance abusers and controls. *Drug and alcohol dependence, 71*(2), 133–141. https://doi.org/10.1016/s0376-8716(03)00090-5

Rixen, J., Meinhardt, L., Glöckler, M., Ziegenbein, M.-L., Schlothauer, A., Colley, M., Rukzio, E., & Gugenheimer, J. (2023). The Loop and Reasons to Break It: Investigating Infinite Scrolling Behaviour in Social Media Applications and Reasons to Stop. *Proceedings of the ACM on Human-Computer Interaction, 7,* 1-22. https://doi.org/10.1145/3604275

Roelofsma, P. H. M. P., & Read, D. (2000). Intransitive intertemporal choice. *Journal of Behavioral Decision Making, 13*(2), 161-177.

Ruhm, C. J. (2012). Understanding overeating and obesity. *Journal of Health Economics, 31*(6), 781-796. https://doi.org/10.1016/j.jhealeco.2012.07.004

Saffer, H. (1991). Alcohol advertising bans and alcohol abuse: An international perspective. *Journal of Health Economics, 10*(1), 65-79. https://doi.org/10.1016/0167-6296(91)90017-H

Samuelson, P. A. (1937). A Note on Measurement of Utility. *The Review of Economic Studies, 4*(2), 155–161. https://doi.org/10.2307/2967612

Sapolsky, R. M. (2017). *Behave: The biology of humans at our best and worst.* Penguin Books.

Sayette, M. A., Loewenstein, G., Kirchner, T. R., & Travis, T. (2005). Effects of smoking urge on temporal cognition. *Psychology of addictive behaviors, 19*(1), 88–93. https://doi.org/10.1037/0893-164X.19.1.88

Schlinger, H. D., & Blakely, E. K. (1994). The Effects of Delayed Reinforcement and a Response-Produced Auditory Stimulus on the Acquisition of Operant Behavior in Rats. *The Psychological Record, 44*(3), 391-409. https://doi.org/10.1007/BF03395923

Shamosh, N.A., & Gray, J.R. (2008). Delay Discounting and Intelligence: A Meta-Analysis. *Intelligence, 36,* 289-305.

Shizgal, P., & Conover, K. (1996). On the Neural Computation of Utility. *Current Directions in Psychological Science, 5*(2), 37–43. http://www.jstor.org/stable/20182387

Simon, H.A. (1990). Bounded Rationality. In: Eatwell, J., Milgate, M., Newman, P. (eds) Utility and Probability. The New Palgrave. Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-20568-4_5

Sozou, P. D. (1998). On Hyperbolic Discounting and Uncertain Hazard Rates. *Proceedings: Biological Sciences*, *265*(1409), 2015–2020. http://www.jstor.org/stable/51159

Steffen, P. R., Hedges, D., & Matheson, R. (2022). The Brain Is Adaptive Not Triune: How the Brain Responds to Threat, Challenge, and Change. *Frontiers in psychiatry, 13,* 802606. https://doi.org/10.3389/fpsyt.2022.802606

Story, G. W., Vlaev, I., Seymour, B., Darzi, A., & Dolan, R. J. (2014). Does temporal discounting explain unhealthy behavior? A systematic review and reinforcement

learning perspective. *Frontiers in behavioral neuroscience, 8*, 76. https://doi.org/10.3389/fnbeh.2014.00076

Strotz, R. H. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies, 23*(3), 165–180. https://doi.org/10.2307/2295722

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). The MIT Press.

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

Thaler, R.H. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8,* 201-207.

Thaler, R. H., & Benartzi, S. (2004). Save More TomorrowTM: Using Behavioral Economics to Increase Employee Saving. *Journal of Political Economy, 112*(S1), S164–S187. https://doi.org/10.1086/380085

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

Tsukayama, E., & Duckworth, A. L. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making, 5*(2), 72–82. doi:10.1017/S1930297500000930

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

Urminsky, O. & Zauberman, G. (2016). The Psychology of Intertemporal Preferences. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making*. John Wiley & Sons,

Volpicelli, J. R., Clay, K. L., Watson, N. T., & Volpicelli, L. A. (1994). Naltrexone and the Treatment of Alcohol Dependence. *Alcohol health and research world, 18*(4), 272–278.

Vuchinich, R. E., & Heather, N. (2003). Introduction: Overview of Behavioural Economic Perspectives on Substance Use and Addiction. In R. E. Vuchinich & N.

Heather (Eds.), *Choice, Behavioural Economics and Addiction* (pp. 1-31). Pergamon. https://doi.org/10.1016/B978-008044056-9/50042-7

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.