Intermittent Discounting

Alexis Direr

Intermittent Discounting

Alexis Direr *

July 19, 2020

Abstract

A novel theory of time discounting is proposed in which future consumption is less valuable than present consumption because of waiting costs. Waiting is intermittent as individuals' attention is periodically distracted away from future gratifications. The more individuals expect to pay attention to the reward, the more they are impatient. The model revisits the fundamental link between short and long-term impatience and solves two behavioral anomalies: impatience over short durations and sub-additive discounting.

J.E.L. codes: D8, E21

Keywords : time preferences, short-term impatience, sub-additive discounting.

*Univ. Orléans, CNRS, LÉO, FRE 2014. Address: rue de Blois - BP 26739, 45067 Orléans Cedex 2 France. Email: alexis.direr@univ-orleans.fr. I thank for useful comments Daniel Heyen, Jean-Pascal Gayant, Louis Raffestin, Maxime Menuet, and participants of the 2017 European Congress of the Econometric Society, the 2018 AFSE Conference, and the seminars of Laboratoire d'Economie d'Orléans, University of Le Mans, and of the Chair of Integrative Risk Management and Economics at ETH Zurich.

1 Introduction

Impatience is a key feature of intertemporal decisions. It is also a versatile property. People do not like waiting two minutes at a stoplight, but are willing to save for their retirement occurring in several decades. These contrasting attitudes suggest that the horizon of choice interacts non-linearly with individuals' propensity to discount future outcomes. This article explores this possibility and proposes a novel theory of time discounting which starts from the observation that waiting for a reward requires a mental effort to resist temptation and cope with some amount of frustration. The more delayed the gratification, the longer the waiting period and the less valuable future utility. Moreover, people do not permanently experience waiting as they spend most of their time absorbed in daily activities during which future gratifications are not reminded. Waiting episodes can be triggered by an external event or a cue, like discussing a new model of cell-phone with a colleague, watching a tv advertisement or contemplating a piece of chocolate fudge cake at a friend's birthday.[1] Reminding may also spontaneously occur when the image of a gratification springs to mind, or when a need is felt, out of boredom, discomfort, stress, hunger, thirst or craving.

To investigate implications of intermittent waiting for time preferences, I pose a multi-period setting in which an agent derives utility from a good which may be consumed now or later. Waiting is both costly and intermittent, as reminding future consumption occurs with some probability every period. When individuals experience intermittent reminding, both the waiting costs and frequency of waiting periods undermine consumer's willingness to delay consumption. Whereas most models of discounting focus on the extent to which people discount future

[1] The frequency of reminders may be reinforced by biased attention toward temptation cues. For example, smokers have been found to display selective attention for smoking-related cues (Mogg, Field, and De Houwer, 2003), and heavy drinkers toward alcohol-related cues (e.g. Townshend and Duka, 2001). See Bernheim and Rangel (2004) for a theoretical analysis.

utilities, implications of intermittent waiting have not been investigated so far. This second dimension of discounting is arguably as important as the size of discounting. For example, a typical question asked to people who suffer from addiction is "how many times a day do you think about ...". In less extreme situations, repeated exposure to temptation goods may lead consumers to indulge, which is routinely exploited by the advertising industry.

I then use the model to investigate the under-explored yet important issue of how short-term discounting is connected to long-term discounting. Several studies have shown that people tend to express strong impatience over short periods (Frederick, Loewenstein and O'Donoghue, 2002). Even small deviations from perfect patience over short durations lead to excessive impatience over long durations due to the power of compounding (Rabin, 2002; Shapiro, 2005; O'Donoghue and Rabin, 2015). Another branch of the literature has reported in experiments sub-additive discounting, i.e. people discount more future payoffs in a sequence of short-duration trade-offs than in a single trade-off over the same time interval (Read, 2001, Read and Roelofsma, 2003, Scholten and Read, 2006 and Kinari et al., 2009). Dohmen et al., (2012) and Dohmen et al. (2017) also find sub-additivity in German representative samples. Present bias models provide only a partial explanation for excessive short-term impatience. Even after controlling for the possibility of present bias, Andreoni and Sprenger (2012) estimate annualized discount rates between 25 and 38 percent and Balakrishnan, Haushofer and Jakiela (2020) between 77 and 96 percent. These values are far in excess than those obtained from introspection or routinely used in macroeconomic models. Moreover, present bias models belong to the class of additive models and therefore cannot account for sub-additivity.

The intermittent discounting model proposes a theory of short-term impatience and sub-additivity based on the premise that individuals expect relatively more waiting and therefore are more impatient in short-delay trade-offs than in a long-delay trade-off during which they expect to spend most of their time forgetting the reward. I also show that when reminding probabilities are stationary and waiting costs are exponentially discounted, the discount function takes a simple two-parameter functional form $D(0) = 1$ and $D(t) = \pi\beta^t + (1-\pi)\beta$, with π the probability of reminding consumption at any future date and $\beta \in (0, 1)$ an inverse measure of waiting costs. The formulation distinguishes two dimensions of impatience: the waiting intensity and its frequency. It boils down to the exponential model of Samuelson (1937) when reminding repeats every period ($\pi = 1$). Exponential discounters are strongly impatient in terms of reminding frequency.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 explains why previous models of discounting have difficulties in properly relating short and long-duration impatience. Section 4 lays out an axiomatically founded model of consumption with intermittent waiting. Section 5 presents an analytically tractable two-parameter version of the model. The two next sections explain why intermittent waiting is consistent with impatience over short durations (Section 6) and sub-additive discounting (Section 7). Section 8 concludes.

2 Related Literature

The paper is related to the vast literature in psychology on waiting, distractions, and time perception. A body of consistent evidence shows that the perception of duration is affected by attention. The father of American psychology William James already noted in 1890: "The tracts of time (...) shorten in passing whenever we are so fully occupied with their content as not to note the actual time itself. (...) On the contrary, a day full of waiting, of unsatisfied desire for change, will seem a small eternity". Closer to us, experimental evidence shows that the ratio of judged to real duration increases when attention is stimulated.[2] People who are paying attention to time itself, e.g. when they are waiting in a queue, or when they have been told in advance to estimate a period of time, feel the time passing more slowly. On the contrary, the ratio of judged to real time decreases when subjects are kept busy by a cognitively demanding task (Zakay and Block, 1997). If attention is distracted by non-temporal information, less capacity is available for processing temporal information (Kahneman, 1973). Katz, Larson and Larson (1991) find that distractions like watching a news board or television while waiting make the wait more acceptable for customers. The evidence is consistent with the model's assumption that consumers pay attention to time in waiting states. The process of waiting causes a lengthening of the perceived temporal distance, which deepens the discount on delayed utility.

Existing models of discounting have difficulties in explaining why waiting is more aversive when the reward is physically close, visible, or can be examined. As already noted by Senior (1836)[3], resisting is particularly difficult when one is in the "actual presence of the immediate object of desire." In the famous "marshmallow experiment" by Mischel and Ebbesen (1970) and Mischel, Ebbesen and Raskoff Zeiss (1972), pre-school children were given the choice between one treat immediately or two if they waited for a short period. They found that children waited much longer for a preferred reward when they were distracted from it than when they attended to them directly. When the rewards were out of sight, 75% of children were able to wait the full time (15 minutes). When it was exposed, the mean delay time was only about 1 minute. Successful children developed

[2]See Fraisse (1963) and Thomas and Brown (1974) for evidence. Hicks, Miller and Kinsbourne (1976) and Thomas and Weaver (1975) provide an attention-based theory of this phenomenon. Another interpretation is that people use a subjective internal timer which is slowed down when they are kept busy (Taatgen, Hedderik and Anderson, 2007).

[3]Quoted in Frederick, Loewenstein and O'Donoghue (2002).

strategies of diversion like singing songs or thinking aloud. Mischel, Ebbesen and Raskoff Zeiss (1972) conclude that "attentional and cognitive mechanisms which enhanced the salience of the rewards shortened the length of voluntary delay, while distractions from the rewards, overtly or cognitively, facilitated delay." Multiple follow-up studies have confirmed that keeping in mind the reward hinders the ability to control one-self (Metcalfe and Mischel, 1999).

More recently, Hofmann et al. (2012) investigate with an experience sampling method how often desires in everyday life, like eating, sleeping or drinking, are felt and how often they are enacted or inhibited. They find that people who were the best at self-control reported fewer episodes of temptation rather than better ability to resist temptations. Ent et al. (2015) also show that self-control is linked to avoiding, rather than merely resisting temptation. Traditional theories of intertemporal choice have difficulties in accounting for those observations as pure time preferences are not distinguished from the frequency of temptations.

Relatedly, some researchers argue that decreasing impatience reflects non-linear perception of time. Ebert and Prelec (2007) report that making people pay more attention to the time dimension of the choice (e.g. by letting people focus on the arrival date of an item) has the effect of increasing discounting of the far future. Zauberman et al. (2009) find that making duration more salient to participants lead them to be more sensitive to time horizon, resulting in less similar preference between short and long time horizons.[4]

[4] See also Radu et al. (2011).

3 Relation between short and long-run rates

This section presents two related behavioral anomalies: impatience over short delays and sub-additive discounting.

3.1 Short-run impatience

In experiments, people tend to express strong impatience over short durations (Frederick, Loewenstein and O'Donoghue, 2002). Short-run impatience is hard to reconcile with reasonable long-run impatience. To understand why, consider an exponential discounter whose discount rate and discount factor over a short period of time (e.g. a day or a week) are $\rho \geq 0$ and $D(1) = (1+\rho)^{-1}$ respectively. Compounded over a full year, the psychological long-run rate is $R = (1+\rho)^{-t} - 1$, with t the number of unit periods in a year. Small levels of short-term impatience translate into potentially extreme degrees of impatience. For instance, a tiny discount rate of $\rho = 0.1$ percent over one day leads to a strong annualized discount rate of 44 percent. Such value seems incompatible with individuals engaging in profitable long-term investments like saving for their long term standard of living.

More reasonable levels of long-term impatience compatible with short-term impatience are obtained by including a bias for the present. This can be done with the quasi-hyperbolic model (Phelps and Pollak, 1968, Laibson, 1997) where future utility is exponentially discounted ($D(t) = (1 + \rho)^{-t}$) and an extra weight $D(0) = 1/\alpha > 1$ applies to present utility. The property is interpreted as a consequence of and a validity test for present bias (Rabin, 2002; Shapiro, 2005; O'Donoghue and Rabin, 2015). Hyperbolic models of discounting (e.g. Harvey, 1986, Mazur, 1987, Loewenstein and Prelec, 1992, Bleichrodt, Rohde and Wakker, 2009 and Ebert and Prelec, 2007) display decreasing impatience and are also consistent with short-run impatience as a result.

However these models do not entirely close the gap between micro and macro estimates of time discounting. Even after controlling for present bias, Andreoni and Sprenger (2012) estimate average annualized discount rates between 25 and 38 percent and Balakrishnan, Haushofer and Jakiela (2020) between 77 and 96 percent. These values contradicts introspection and are at odds with discount rates used in macroeconomic models.

Moreover, in the quasi-hyperbolic model, the DM behaves like an exponential discounter in intertemporal trade-offs which do not involve an immediate consumption. Hence any departure from perfect patience in delayed trade-offs over short time intervals like a day or a week leads to implausible long-term impatience, as in the exponential model. For instance, assume the DM is indifferent between consuming 1 at date s and $y > 1$ at date $k > s$, where k is temporally close to s. With quasi-hyperbolic preferences: $\alpha(1+\rho)^{-s}u(1) = \alpha(1+\rho)^{-k}u(y)$, which is equivalent to being indifferent between 1 now and y in k periods in the exponential model: $u(1) = (1+\rho)^{-k}u(y)$. We are back to the quantitative impossibility and 'compounding curse'.

3.2 Subadditivity

Let us define a dated consumption (y, t) with y the quantity consumed and t the consumption date. The immediate quantity x is determined by the indifference condition $(x, 0) \sim (y, t)$. By transitivity of indifference, there exists a payoff z such that $(x, 0) \sim (z, s) \sim (y, t)$. With a discounted utility formulation and the notation $\phi(0, t) = D(t)/D(0)$, the long-term discount factor can be decomposed into the product of two discount factors over shorter time intervals:

$$\phi(0, t) = \phi(0, s)\phi(s, t) \tag{1}$$

The relation holds for all usual time-separable discount functions $D(t)$. However, people tend to be more impatient when confronted with multiple short-delay trade-offs in a sub-divided interval than with a single trade-off over the whole interval (Read, 2001, Read and Roelofsma, 2003, Scholten and Read (2006), Dohmen et al., 2012, Dohmen et al. 2017):

$$\phi(0,t) > \phi(0,s)\phi(s,t) \tag{2}$$

Sub-additivity and its opposite, super-additivity, are defined in accordance:

Definition 1 *For all $x, y, y', z \in X$ and $s, t \in T$, $0 < s < t$, such that $(x,0) \sim (y',t)$, $(x,0) \sim (z,s)$ and $(z,s) \sim (y,t)$. Preferences are additive if $y = y'$, sub-additive if $y > y'$ and super-additive $y < y'$.*

A higher payoff y signals more impatience over repeated short delays than over long horizons. Moreover, the relation (2) can be expanded by subdividing further the whole interval, assuming that the behavioral pattern repeats over smaller intervals:

$$\begin{aligned}\phi(0,t) &< \phi(0,s)\,\phi(s,t) < \phi(0,s-1)\,\phi(s-1,s)\,\phi(s,t)\\ &< \phi(0,s-2)\,\phi(s-2,s-1)\,\phi(s-1,s)\,\phi(s,t)\\ &< ...\\ &< \phi(0,1)\,\phi(1,2)\,...\,\phi(t-2,t-1)\,\phi(t-1,t)\end{aligned}$$

Because the discount factor gaps between every interval and its subdivisions compound, the long-run discount factor can potentially be significantly lower than the product of short-term factors.

To sum up, short-run impatience and sub-additive discounting share a similar property: short-duration discount rates are 'too large' to square with long-duration discount rates. Hyperbolic or quasi-hyperbolic models of discounting are consistent with short-run impatience but have difficulty in explaining impatience over delayed short durations and cannot account for sub-additive discounting. The rest of the paper proposes a unified theory of short-run impatience, and sub-additivity based on intermittent waiting costs.

4 Time preference with waiting

4.1 Expected utility with waiting

I proceed in two steps. I first pose a general expected utility model of consumption and random waiting before presenting a full-fledged model of intertemporal choice. A consumer decides at which date $t \in T = \{0, ..., \bar{t}\}$ a good, which quantity is $x \in X = (0, \bar{x}]$, is consumed. Let $\theta = (\theta_s, s = 0, ..., \bar{t}) \in \Theta = \{0, 1\}^{\bar{t}+1}$ be an exogenous temporal sequence of reminding. $\theta_s = 1$ if the DM pays attention to the consumption good in period s. A reminding period is a waiting period if the good has not already been consumed in the past and is not consumed during the current period. $\theta_s = 0$ if s is not a reminding period and consequently not a waiting period. The objects of choice are lotteries with finite support. Define the set of lotteries

$$L = \left\{ P : (X, T, \Theta) \to [0, 1] \; ; \; \sum_{\Theta} P(x, t, \theta) = 1 \right\}$$

Choices expressed at time 0 are modeled by a binary relation $\succcurlyeq$ on L.[5] Define the mixing operation of lotteries P and Q with $a \in [0, 1]$, as $aP + (1 - a)Q = a \sum_{\Theta} P(x, t, \theta) + (1 - a) \sum_{\Theta} Q(x', t', \theta)$. The vNM axioms are:

A1. Weak order $\succcurlyeq$ is complete and transitive.

[5] The issue of time-consistency is sidestepped by focusing on time preferences from date 0 perspective, as if the DM could commit to them.

A2. Continuity For all P, Q, $R \in L$, if $P \succ Q \succ R$, there exist a, $b \in (0,1)$ such that

$$aP + (1-a)R \succ Q \succ bP + (1-b)R$$

A3 Independence For all P, Q, $R \in L$, and $a \in (0,1)$, $P \succcurlyeq Q$ iff

$$aP + (1-a)R \succcurlyeq aQ + (1-a)R$$

The DM behaves as if she is maximizing the expected value of some function u_{vNM} defined over the potential outcomes.

Theorem 1 (vNM Preference) $\succcurlyeq$ satisfies A1-A3 if and only if there exists $u_{vNM} : (X, T, \Theta) \to \mathbb{R}$ such that, for every P, $Q \in L$, $P \succcurlyeq Q$ iff

$$\sum_{\Theta} P(x,t,\theta) u_{vNM}(x,t,\theta) \geq \sum_{\Theta} Q(x',t',\theta) u_{vNM}(x',t',\theta)$$

The proof is not reproduced and can be found for instance in Mas-Colell, Whinston, and Green (1995, pp. 176-178). The model, which does not impose at this stage meaningful properties on the utility function, is specialized in the next section.

4.2 Time preferences

We now put more structure on admissible lotteries and time preferences. Given lottery P, probability $p_s \in [0,1]$ of reminding the reward in period $s = 0, 1, ..., T$ is defined by $p_s = \sum_{\Theta} P(x,t,\theta \,|\, \theta_s = 1)$. The sequence of dated reminding probabilities is denoted $p = (p_0, p_1, ..., p_{\bar{t}}) \in [0,1]^{\bar{t}+1}$. For convenience, time preferences are now defined by $\succcurlyeq'$ over dated consumptions and reminding probabilities $(x,t,p) \in X \times T \times [0,1]^{\bar{t}+1}$. The relation entails the same ordering as $\succcurlyeq$: $P \succcurlyeq Q \iff (x,t;p) \succcurlyeq' (y,s;q)$ with $p_r = \sum_{\Theta} P(x,t,\theta \,|\, \theta_r = 1)$ and $q_r = \sum_{\Theta} Q(y,s,\theta \,|\, \theta_r = 1)$ for all $r \in T$.

Time preferences are restricted by the following axioms. Axiom A4 ensures that the good is valuable to the DM for any sequence of reminding probabilities.

A4. Monotonicity For all $x, y \in X$, $y > x$, $t \in T$ and $p \in [0,1]^{\bar{t}+1}$, $(y, t; p) \succ' (x, t; p)$.

The next axiom is an adaptation of the Thomsen condition. It ensures the multiplicative separability of the discount factor and the utility function and is weaker than the stationarity axiom (Fishburn and Rubinstein, 1982).

A5. Thomsen separability For all $x, y, z \in X$, $t, s, r \in T$ and $p \in [0,1]^{\bar{t}+1}$, $(x, t; p) \sim' (y, s; p)$ and $(z, t; p) \sim' (y, r; p) \Longrightarrow (x, r; p) \sim' (z, s; p)$.

With words, if $y - x$ is needed to compensate for the additional delay of $s - t$, and $z - y$ for the additional delay of $t - r$, then $(y - x) + (z - y) = z - x$ is needed to compensate for the additional delay of $(s - t) + (t - r) = s - r$.

According to A6, a preference between two dated consumptions with a common reminding probability is not affected by variations of this probability. The axiom ensures that waiting costs are time additively separable.

A6. Time additivity For all $x \in X$, $t, r \in T$ and $p, p', q, q' \in [0,1]^{T+1}$ such that $p' = p$ and $q' = q$ except $p'_r \neq p_r$ and $q'_r \neq q_r$, if $p_r = q_r$ then $(x, t; p) \succcurlyeq' (x, t; q) \Longrightarrow (x, t; p') \succcurlyeq' (x, t; q')$ for all $p'_r = q'_r \in [0,1]$.

The DM is 'wait-averse'. She prefers waiting to be less likely all else equal:

A7. Waiting aversion For all $x \in X$, $t \in T$ and $p \in [0,1]^{\bar{t}+1}$, $p' = p$ except $p'_s > p_s$, $(x, t; p) \succ' (x, t; p')$ if $\forall s < t$ and $(x, t; p) \sim' (x, t; p')$ if $\forall s \geq t$.

By definition of waiting, reminding is costly before but not during or after consumption. The next axiom is a key axiom for the theory. It makes the DM time neutral in absence of reminding.

A8. Temporal indifference For all $x \in X$, $t \in T - \{\bar{t}\}$ and $p \in [0,1]^{\bar{t}+1}$,

$(x, t; p) \sim' (x,\, t+1; p)$ if $p_t = 0$.

A8 states that if date t cannot be a reminding period, the DM is indifferent between consuming at this period or next one. It formalizes the intuition that people may delay consumption effortlessly if they are distracted by unrelated activities. For instance the DM may be willing to postpone watching the last James Bond until evening if she expects no to remind the movie during the day. By extension, if periods t to $t + k$ have zero probability of reminding, the DM is indifferent between consuming at dates t, $t + k + 1$, or any date within the time interval.

Axioms A7-A8 define together a weak form of impatience. They are equivalent to A7-A8b with:[6]

A8b. Weak impatience For all $x \in X$, $t \in T - \{\bar{t}\}$ and $p \in [0, 1]^{\bar{t}+1}$, $(x, t; p) \succ (x,\, t + 1; p)$ iff $p_t > 0$.

As in traditional models of time discounting, an impatient DM prefers consuming the earliest period. She is only weakly impatient here as impatience is conditional on the good to be recalled to mind with a strictly positive probability.

While waiting is only felt during reminding periods, consumption is identically valued regardless whether the period is a reminding or non-reminding state:

A9. Static indifference For all $x \in X$, $t \in T$, $p \in [0, 1]^{\bar{t}+1}$, $p' = p$ except $p'_t \neq p_t$, $(x, t; p) \sim' (x, t; p')$ for all $p'_t \in [0, 1]$.

[6] $(x, t; p) \sim (x, t + 1; p)$ if $p_t = 0$ (A7) and $(x, t + 1; p) \succ (x, t + 1; q)$ with $q = p$ except $q_t > p_t = 0$ (A8). Obviously $(x, t + 1; p) \sim (x, t + 1; q)$ if $q_t = p_t$.

4.3 Utility representation

Axioms A1-A9 are consistent with a theory of discounted expected utility with waiting costs (see Proof in Appendix):

Proposition 1 *$\succcurlyeq$' satisfies A1-A9 if there exists $u : X \to \mathbb{R}_+$ such that, for all (x, t, p), $(y, s, q) \in X \times T \times [0, 1]^{\bar{t}+1}$, $(x, t, p) \succcurlyeq (y, s, p)$ if*

$$-\sum_{j=0}^{t-1} p_j \delta_j u(x) + u(x) \geq -\sum_{j=0}^{s-1} q_j \delta_j u(y) + u(y)$$

with δ_j, $j = 0, 1, ..., T$, satisfying

$$1 > 1 - \delta_0 > 1 - \delta_0 - \delta_1 > ... > 1 - \delta_0 - ... - \delta_{T-1} > 0 \tag{3}$$

Date 0 intertemporal utility of (x, t, p) is the sum of expected waiting costs $\delta_j u(x)$ accumulated before the good is consumed and utility $u(x)$ from consuming x at date t. When consumption is postponed to date t, the DM may remind the reward and incur waiting costs with probability p_j every period before t. In accordance with A5 (Thomsen separability), the disutility is proportional to deferred utility $u(x)$, with the intuition that the more pleasurable the outcome, the more unpleasant the waiting.

Note also that absent waiting costs, utility of consumption would be valued the same way whatever the consumption date. Likewise, waiting costs depend on the date at which they are incurred, but not on the remaining delay until consumption. Both results stem from A8 (temporal indifference) which states that delaying consumption from date $t-1$ to t is harmless if the DM does not expect to pay attention to the reward at $t-1$ ($p_{t-1} = 0$). The two outcomes $(x, t-1, p)$ and (x, t, p) are identically valued by the DM despite varying delays between the present and the consumption date or between the waiting periods and the consumption date. Since utility is not time discounted per se, impatience entirely rests on anticipated waiting costs.

From A5 (Thomsen separability), the intertemporal utility function can be expressed as the product of a discount function D and period utility u:

$$D(t)u(x) = \Big(1 - \sum_{j=0}^{t-1} p_j \delta_j\Big) u(x)$$

with $D(0) = 1$. In accordance with A4 (monotonicity), condition (3) requires that consumption is valuable at every horizon, i.e. $D(t) > 0$ for all $t \in T$, even in the most unfavorable environment in which the DM waits every period before consuming (all $p_j = 1$). (3) also implies that the longer the delay before consuming, the smaller the sum of temporal weights attached to utility: $D(0) \geq D(1) \geq D(2) \geq ... \geq D(T) \geq 0$, whatever reminding probabilities. The decrease of the discount factor with delay is the classical definition of impatience. The further away consumption is delayed, the greater the number of periods during which the DM may remind future consumption and the lower expected utility. The decrease is non-linear however, as she may expect (possibly long) periods during which consumption is not recalled.

Proposition 1 should be valid for an arbitrarily large number of periods, especially when the unit of time is short, like a day or an hour. Let us define the asymptotic minimal utility $D_{min}u(x)$ as the infinitely postponed discounted utility with maximal waiting costs with all probabilities set to 1:

$$D_{min}u(x) = \lim_{T \to \infty} (1 - \delta_0 - \delta_1 - ... - \delta_{T-1}) u(x) \tag{4}$$

Condition $D_{min}u(x) \geq 0$ extends condition (3) to the infinite horizon case. The condition implies that temporal weights δ_s become arbitrarily close to each other as the sequence progresses.[7]

[7] Using the fact that any convergent sequence is a Cauchy sequence, for any given $\varepsilon > 0$, there exists a date T_0 such that for any pair of dates (s, t) satisfying $T_0 < s < t$, we have $|D(t) - D(s)| < \varepsilon$ or $\delta_s + \delta_{s+1} + ... + \delta_{t-1} < \varepsilon$.

5 A two parameter discounting model

This section presents a simplified and more tractable model of discounting. Although the results of the next sections do not rely on the simplified model, it is worth presenting for three reasons. First, it makes clear how interact the two dimensions of discounting: the waiting costs and the frequency with which consumption is reminded. Second, the model nests the canonical exponential model as a special case and allows interesting cross-interpretations. Third, it reduces the number of free parameter to two, which eases the calibration of the model. It is obtained by adding two axioms and two assumptions.

First, the DM prefers waiting as late as possible.

A10. (Preference for late waiting) For all $x \in X$, $t \in T$ and $p \in [0,1]^{\bar{t}+1}$, $(x,t;p) \succ (x,t;q)$ if $q = p$ except $q_{j+1} = p_j < p_{j+1} = q_j$ for all $j < t-2$.

The DM prefers to swap two temporally adjacent reminding probabilities if it results in delaying the higher probability. The axiom is supported by the common observation that people tend to postpone unpleasant feelings or tasks. Note that the alternative property, not explored further in the paper, would be consistent with the DM experiencing craving for the good. Early waiting would be preferred in this case as the feeling of deprivation is expected to build up.

The second axiom allows a smooth evolution of waiting costs with delay. It states that DM's preference relative to the timing of waiting evolves smoothly with delay.

A11. (Preference smoothness) For all $x \in X$, $t \in T$ and $p \in [0,1]^{\bar{t}+1}$, there exists $\beta > 0$ such that $(x,t,p) \sim (x,t;q)$, $q = p$ except $q_j = p_j + \Delta$ and $q_{j+1} = p_{j+1} - \dfrac{\Delta}{\beta}$ for all $\Delta \in (0, min(1-p_j, \beta p_{j+1}))$ and $j < t-2$.

An increase of the waiting probability at date j by the margin Δ leaves the

DM indifferent if one period later, the waiting probability is decreased by the margin Δ/β, where β is a common coefficient for all dates.

A10 and A11 makes explicit the discounting of waiting costs (see Proof in Appendix):

Proposition 2 *Under A1-A11, there exists* $\beta \in (0,1)$ *such that* $\delta_t = \beta^t \delta_0$*, with* $\delta_0 < \dfrac{1-\beta}{1-\beta^t}$.

Waiting costs are discounted by the exponential factor β^t when A11 is added. A10 (preference for late waiting) restricts $\beta < 1$. Condition $\delta_0 < \dfrac{1-\beta}{1-\beta^t}$ ensures that utility delayed arbitrarily far in the future, net of waiting costs, remains positive.

Next, minimal utility (4) with infinite horizon is normalized to zero: $D_{min}u(x) = 0$ for all $x \in X$. The assumption is consistent with the requirement that infinitely delayed utility is useless.[8] Together with A1-A11, it implies the parameter restriction:[9]

Assumption 1 $\delta_0 = 1 - \beta$.

The higher the immediate waiting costs δ_0, the heavier future waiting costs must be discounted so as intertemporal utility remains non-negative.

Last, all periods but the present have a common probability of reminding.

Assumption 2 $p_j = \pi \in [0,1]$*, for all* $j \in T$*, except* $p_0 = 1$.

---

[8] $\lim_{t\to\infty} D(t)$ is satisfied by all common models of intertemporal choice, including the exponential specification and the hyperbolic and quasi-hyperbolic models. With generalized hyperbolic preferences: $\lim_{t\to\infty}(1+ht)^{-r/h}u(x) = 0$, with $h, r > 0$. With quasi-hyperbolic discounting: $\lim_{t\to\infty} \alpha\beta^t u(x) = 0$, given $0 < \alpha, \beta < 1$.

[9] Since $\lim_{t\to\infty} D(t) = 1 - \delta_0 - \beta\delta_0 - \beta^2\delta_0 - ... - \beta^{t-1}\delta_0 = 1 - \delta_0/(1-\beta) = 0$.

The present is a special date. Conditional on not consuming the good, it is either a waiting period *with certainty* or a forgetful period. Examples of present reminding periods are decision or planning dates. Even if consumption is not yet available, choosing between alternative plans may act as a cue to consume and trigger waiting costs. Many present periods are certainly not decision periods, but, as a practical matter, we only need to specify DM's preferences when in a decision or reminding state.

The additional axioms and assumptions lead to a simple and intuitive discount function (see Proof in Appendix):

Proposition 3 *Under A1-A11 and Assumptions (1)-(2),* $D(0) = 1$ *and*

$$D(t) = \pi\beta^t + (1-\pi)\beta \tag{5}$$

with $\beta \in (0, 1)$ *and* $\pi \in [0, 1]$ *a constant probability of reminding.*

The parameter β is the rate used to discount future waiting costs in Proposition 2 but is also an inverse measure of waiting costs (Ass. 1). The discount factor $D(t)$ is a probability-weighted mean of two discount functions in which the reminding frequency π plays a pivotal role.[10] The smaller π, the more patient the DM. Patience is maximal if the DM does not expect to remind the reward in the future ($\pi = 0$), implying a constant discount $D(t) = \beta$ after date 0.

To the opposite, if the DM is so impatient that she expects to remind the reward every period ($\pi = 1$), discounting is exponential. Preferences inherit the normative features of the exponential model: constant impatience and time

[10] The duality may also be interpreted as reflecting the conflict of two selves or systems. One self is impatient and discounts exponentially. The second is more patient and equally discounts all future utilities. The higher the reminding probability, the greater the weight given to the impatient self. McClure et al. (2007) offer a similar interpretation for the quasi-hyperbolic model. See also Ainslie (1992) and Metcalfe and Mischel (1999).

consistency. However, when not only the discount factor β but also the waiting frequency are considered, exponential discounting signals an extreme form of impatience. This is is consistent with the fact that even small deviations from perfect patience make discounted utility rapidly converging to zero with the passage of time. From a normative perspective, the DM would like to minimize waiting costs by avoiding reminding future consumption too often, which an exponential discounter fails to do.

At a more fundamental level, the reminding probability π reflects individuals' attitude toward time. Individuals with a small π choose as if temporally distant payoffs were close to the present. Opting for a larger and late reward does not look like a high sacrifice. Conversely, individuals with a high π behave as if temporally close payoffs were delayed far in the future. It is as if time is stretching in the first case and accelerating in the second case. In extreme situations of substance addiction, a high reminding frequency coupled with a hourly or daily time frame make individuals strongly impatient. Drug abusers demonstrate shortened time horizons and decreased sensitivity to delayed consequences (Petry, Bickel and Arnett, 1998). Smokers (Baker, Johnson and Bickel, 2003), alcoholics (Vuchinich and Simpson, 1998) or substance-dependent individuals (Kirby, Petry, and Bickel, 1999) show large discounting of delayed rewards.

Having laid the foundations of the wait-based model, we now turn to its behavioral implications.

6 Short-delay impatience

The wait-based model of discounting is consistent with short-delay impatience described in Section 3 and proposes a behaviorally founded interpretation of the property. Two types of trade-offs are distinguished: those with an immediate payoff and those in which all alternatives are in the future.

6.1 Impatience with an immediate payoff

Under A1-A9, DM's short-term preferences are given by $D(1)/(D(0) = 1 - p_0\delta_0$, whereas the long-term discounting function is:

$$D(t)/(D(0) = 1 - p_0\delta_0 - p_1\delta_1 ... - p_{t-1}\delta_{t-1}$$

Significant impatience over short-delays (as measured by $D(1)^{-1} - 1$) and moderate impatience over long delays (as measured by $D(t)^{-1} - 1$) can be jointly obtained if (i) the present is a reminding period ($p_0 = 1$) and (ii) subsequent episodes of reminding are infrequent, i.e. p_s are small for all $s > 0$. Condition (i) is consistent with Assumption 2 and is motivated by the immediate availability of the reward. It is consistent with a bias for present consumption since waiting is felt in the present as opposed to anticipated with some probability in the future. Condition (ii) is realistic given that most individuals spend a small fraction of their time thinking about future consumption. Long-term impatience remains bounded as a result, even though present waiting costs δ_0 are large.

The relation between short-run and long-run discounting can be investigated further by using the simplified model with exponentially discounted waiting costs (see Proposition 3): $D(t) = \pi\beta^t + (1-\pi)\beta$. If $\pi = 1$, reminding repeats over and over, with the consequence that present and future periods look alike. The symmetry leads to the exponential model and its inability to plausibly account for both short-run and long-run impatience. If reminding is infrequent, the discount factor is close to β whatever time horizon t. The actual frequency of reminding is therefore a key factor determining to what extent short-run impatience translates into long-run impatience.

To get a quantitative assessment of the relationship between short and long-run impatience, let us start wit a daily time unit. With costly waiting, the implicit psychological short-term rate ρ is defined by $(1+\rho)^{-1} = D(1) = \beta$. The implicit long-run rate R is defined by $(1+R)^{-1} = D(365) = \pi\beta^{365} + (1-\pi)\beta$. Table 1 shows long-run rates R for various values of short-run rates ρ and reminding probabilities π.

Table 1: Implicit long-run rate R (in percent) in function of the short-run rate ρ (in percent) and reminding probability π

	$\pi = 1$	0.5	0.3	0.1	0.05	0.01	0
$\rho = 0.1$	44	18.1	10.2	3.2	1.65	0.41	0.10
1	3,678	96.7	42.7	11.9	6.17	1.99	1
2	1.37×10^5	103.8	45.7	13.3	7.36	3.03	2
5	5.42×10^9	110.0	50.0	16.7	10.5	6.06	5
10	1.28×10^{17}	120.0	57.1	22.2	15.8	11.1	10

The lower the reminding probability, the closer the long-run rate to the short-run rate. The limit case $\pi = 1$ corresponds to the present-neutral exponential model in which long-run rates take implausible high values. To the contrary, the intermittent wait-based model is able to account for both non-trivial short-term impatience and reasonable long-run impatience. Even for daily short-run rates as large as 10%, the long-run rate is only 5 percentage points higher when the DM reminds the future reward 5% of the time.

The wait-based model puts forth a fundamental reason why long-run rates do not reduce to a compound of short-run rates. The latter are generally elicited with subject's attention caught and oriented toward a concrete choice in which immediate consumption is feasible. It is therefore not surprising that short-run impatience is not trivial. Yet, extrapolating long-run rates by repeatedly compounding the obtained short-run rate is like presuming that the DM is placed in the same short-term decision situation over and over. Instead, it is more realistic to expect that she forgets the reward most of the time. This provides a strong intuition of why impatience over long-delay trade-offs is likely to remain in a reasonable range.

6.2 Impatience over delayed short-durations

Because present bias models put an extra weight on present utility, they cannot account simultaneously for significant impatience over short time intervals starting in the future and reasonable impatience over long periods (see Section 3). With intermittent waiting, suppose the DM is indifferent between 1 unit in s periods and y units in k periods:

$$\big(1-p_0\delta_0-p_1\delta_1-...-p_{s-1}\delta_{s-1}\big)u(1) = \big(1-p_0\delta_0-...-p_s\delta_s-...-p_{k-1}\delta_{k-1}\big)u(y) \quad (6)$$

In long-term trade-offs, let us define y' such that she is indifferent between 1 unit immediately and y' units at date $t \geq k$:

$$u(1) = \big(1 - p_0\delta_0 - ... - p_s\delta_s - ... - p_t\delta_t\big)u(y') \quad (7)$$

When pondering her choice in the delayed trade-off (6), the date s stands out from other dates. The DM may picture herself at date s and anticipate the waiting costs associated with postponing consumption to date k. Imagining a hypothetical event makes the event seem more likely through the use of the availability heuristic (Carroll, 1978, Kahneman et al., 1982). Moreover, the probability may be objectively high if the DM anticipates that the reward will be ostensibly displayed at date s. It follows that the reminding probability in the first date of the trade-off is likely to be high compared to probabilities in other dates. This will also be the case if the propensity to imagine the future vividly makes delayed goods appear to be temporally closer than they actually are (Becker and Mulligan, 1997). In the model, this effect means that reminding probabilities are low until the consumption date. The argument is formalized in the following assumption with $\longleftrightarrow$ a shortcut for 'is compared to':

Assumption 3 *For all* $z, y, x \in X$ *and* $r, s, k, t \in T$, $0 \leq r < s < k \leq t$, *if* $(z, s, p) \longleftrightarrow (y, k, p)$ *and* $(x, r, p') \longleftrightarrow (y', t, p')$ *then* $p_s > p'_s$.

The time intervals of the two trade-offs overlap and share the common period s. But whereas s is the first date of the trade-off in the first choice, it is an ordinary period in the second one. Assumption 3 can be applied to trade-offs (6) and (7) with $r = 0$ and $z = x = 1$. Although a high reminding probability at date s makes the DM impatient over a short time interval, it does not mean that she is also excessively impatient over long durations if other reminding probabilities p_j, $j \neq s$, are low enough.

To illustrate the point, let us reinterpret the estimates found by Balakrishnan, Haushofer and Jakiela (2020). After controlling for utility curvature and present bias, they estimate an average discount rate between 5.52 and 7.65 percent over two and four-week time intervals. Since the trade-of is delayed by roughly a day (from morning to end of the day), such high discount rates are not explained by present bias. Compounded over a full year, the annualized discount rate is between 327 and 632 percent, which seems implausible. Let us use the simplified version of the model with exponentially discounted waiting costs (see Proposition 2). The DM is indifferent between obtaining 1 unit of money at the end of the day and $y = 1.0765$ units in two weeks. Assuming a daily frequency, Condition

(6) with $s = 1$ and $k = 14$ becomes

$$\beta = \frac{u(y)}{u(1)}\big(\pi\beta^k + (1-\pi)\beta - (p_1 - \pi)\beta(1-\beta)\big)$$

The condition simplifies with sure reminding the first date of the trade-off ($p_1 = 1$) to

$$\beta = \frac{u(y)}{u(1)}\big(\pi\beta^k + (1-\pi)\beta^2\big) \tag{8}$$

The corresponding long-term trade-off offers 1 unit in one period (at the end of the day) and y' units in one year ($t = 365$). The indifference condition is:

$$\beta = \frac{u(y')}{u(1)}\big(\pi\beta^t + (1-\pi)\beta\big) \tag{9}$$

Conditions (8) and (9) can be easily compared. If reminding repeats every period ($\pi = 1$), we are back to the exponential case and the 'compounding curse'. The term β^t rapidly converges to zero with horizon t, leading to a rapidly increasing ratio $u(y')/u(1)$ and inflating impatience. To the contrary, when reminding is unlikely at all future dates but date 1, reasonable impatience over long durations is compatible with departure from perfect patience over short durations.

To continue with the example, a relative risk aversion coefficient equal to 0.619 is assumed, which is the value estimated by by Balakrishnan, Haushofer and Jakiela (2020):

$$\frac{u(y)}{u(1)} = 1.0765^{1-0.619}$$

Suppose that the DM expects to remind the reward 5 percent of the time ($\pi = 0.05$). From the indifference condition (8), we obtain $\beta = 0.982$. Once those values are plunged into condition (8), we find $\frac{u(y')}{u(1)} = 1.0530$ and an annualized rate of return equal to 14.5 percent. The rate is only twice the rate found over a two week time interval and is much lower that the rate of 632 percent computed by the authors.

7 Subadditive discounting

Assume indifference between z in $s > 0$ periods and y in $t > s$ periods. With multiplicatively separable utility function, $D(s)u(z) = D(t)u(y)$. Likewise, the DM is indifferent between x now and y' in t periods if $D(0)u(x) = D(t)u(y')$. If preferences are additive in the sense of Definition 1, assuming the third equivalence $D(0)u(x) = D(s)u(z)$ guarantees $y = y'$. Note that additivity necessitates that in both trade-offs, $u(y)$ and $u(y')$ are discounted by the *same* factor $D(t)$.

Suppose now that discounting is the result of waiting costs. When choosing between dates s and t, the date s of the trade-off is a pivotal date, for the same reasons as was date s in trade-off (6) with a delayed short duration.The DM is induced to think about consuming the reward at date s, which increases the reminding probability. To the contrary, when choosing between dates 0 and t, nothing special happens in period s during which the DM may be distracted by other occupations. The probability p'_s of reminding the good in this case is most likely lower than the probability p_s in the first trade-off. The assumption is consistent with Read (2001)'s interpretation according to whom sub-dividing a delay undermines people propensity to withstand waiting by making them pay more attention to every part of the delay. In other words, Assumption 3 applies with $r = 0$ and $k = t$. It follows that the two discount factors, respectively denoted $D(t)$ and $D'(t)$, although referring to the same delay, may differ in a systematic way:

$$D(t) = 1 - p_0\delta_0 - ... - p_s\delta_s - ... - p_{t-1}\delta_{t-1}$$

$$D'(t) = 1 - p_0\delta_0 - ... - p'_s\delta_s - ... - p_{t-1}\delta_{t-1}$$

with $D(t) < D'(t)$ since $p_t > p'_t$. Wait-based preferences are sub-additive according to Definition 1 (see Proof in Appendix):

Proposition 4 *Under A1-A9 and Assumption 3, preferences are sub-additive.*

The proposition is obtained for common reminding probabilities across trade-offs except date s. Since the DM jointly considers the two trade-offs, her beliefs should be consistent across her options. Intuitively, the DM expects less waiting in a long-term trade-off than in a sequence of short-term trade-offs of same length. It follows that the DM is more willing to delay a reward in the first situation than in the second one.

8 Conclusion

A novel theory of time discounting is proposed in which disutility of waiting explains why future utilities are depreciated. It adopts a nonlinear approach of time, more familiar to psychologists, in which experienced time elapses only when attention to future gratifications is paid. As stressed by Stout (1932): “In general, temporal perception is bound up with the process of attention... What measures the lapse of time is the cumulative effect of the process of attending". When discounting is tied to expected episodes of waiting, long duration discount rates are only weakly connected to short duration discount rates. Compounding and annualization of discount rates, should be used with caution. The unit of time over which choices are made becomes important. The model helps bridge the gap between high short-term impatience found in experiments and low interest rates at the macroeconomic level, a point stressed by Cochrane (2011): “People report astounding discount rates in surveys and experiments, yet still own long-lived assets, houses, and durable goods.”

If perception of durations is elastic, discounting is exposed to manipulation, a possibility explored in experiments by Mischel, Ebbesen and Raskoff Zeiss

(1972) or Ebert and Prelec (2007). A more direct test of the theory would consist in proposing trade-offs between smaller rewards now and a larger ones later. Subjects in the treatment group would be informed to be recalled the reward during the waiting time, for instance by watching a video related to it, by letting the reward in plain sight, or, over longer time intervals, by periodically receiving messages about it. The theory predicts that treated subjects should express more impatience than subjects without recalls.

The paper does not investigate why some people experience more temptation episodes than others. Addiction, weak habits, or genetic predispositions may be part of the story. Some people seem better able to avoid potential conflicts, e.g. by installing adaptive routines (Gillebaart and de Ridder, 2015). To this regard, it would be interesting to design experiments which would separately estimate pure discount rates and propensity of forgetting future rewards.

References

Ainslie G. (1992) Picoeconomics: the strategic interaction of successive motivational states within the person. New York: Cambridge University Press.

Andreoni J. and C. Sprenger (2012) “Estimating time preferences from convex budgets” *American Economic Review*, 102 (7), 3333-56.

Baker F., Johnson M. W. and W. K. Bickel (2003) “Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude” *Journal of Abnormal Psychology* 112 (3), 382-392.

Balakrishnan U., J. Haushofer and P. Jakiela (2020) “How soon is now? Evidence of present bias from convex time budget experiments” *Experimental Economics*, 23 (2), 294.

Becker G. S. and C. B. Mulligan (1997) ”On the endogenous determination of time preference“ *Quarterly Journal of Economics*, 112 (3), 729-758.

Bernheim B. D. and A. Rangel (2004) “Addiction and cue-triggered decision processes” *American Economic Review*, 94(5), 1558-1590.

Bleichrodt H., K. I. M. Rohde and P. P. Wakker (2009) “Non-hyperbolic time inconsistency” *Games and Economic Behavior* 66, 27-38.

Carroll J. S. (1978) ”The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic“ *Journal of Experimental Social Psychology* 14 (1), 88-96.

Cochrane J. H. (2011) “Presidential address: Discount rates” *The Journal of Finance*, 66 (4), 1047-1108.

Dohmen T., Falk A., Huffman D. and U. Sunde (2012) “Interpreting time horizon effects in inter-temporal choice", IZA working paper 6385.

Dohmen T., Falk A., Huffman D. and U Sunde (2017) "The robustness and pervasiveness of sub-additivity in intertemporal choice", working paper.

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

Ent M. R., Baumeister R. F., and D. M. Tice (2015) "Trait self-control and the avoidance of temptation" *Personality and Individual Differences*, 74, 12-15.

Fishburn P. C, and A. Rubinstein (1982) "Time preference" *International Economic Review* 23(3) 677-694.

Fraisse P. (1963) The psychology of time. New York: Harper and Row.

Frederick S., Loewenstein G., and T. O'Donoghue (2002) "Time discounting and time preference: A critical review" *Journal of Economic Literature*, pp. 351-401.

Gillebaart M. and D. T. D. de Ridder (2015) "Effortless self-control: A novel perspective on response conflict strategies in trait self-control", *Social and Personality Psychology Compass* 9 (2), 88-99.

Harvey C. M. (1986) "Value functions for infinite period planning" *Management Science* 32, 1123-39.

Hicks R. E., Miller G. W. and M. Kinsbourne (1976) "Prospective and retrospective judgments of time as a function of amount of information processed" *American Journal of Psychology*, 89, 719-730.

Hofmann W., Baumeister R. F., Förster G. and K. D. Vohs (2012) "Everyday temptations: an experience sampling study of desire, conflict, and self-control" *Journal of Personality and Social Psychology* 102 (6), 318-35.

James W. (1890) The Principles of psychology, Dover Publications 1950, chapter XV.

Kahneman D. (1973) Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman D., Slovic P., and A. Tversky (eds.): 1982, Judgment under uncertainty: Heuristics and biases, Cambridge Univ. Press, Cambridge and New York.

Katz K., Larson B. and R. Larson (1991) “Prescriptions for the waiting-in-line blues: entertain, enlighten and engage” *Sloan Management Review* 32, 44-53.

Kinari Y., Ohtake F. and Y. Tsutsui (2009) “Time discounting: Declining impatience and interval effect” *Journal of Risk and Uncertainty*, 39 (1), 87-112.

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

Laibson D. (1997) “Golden eggs and hyperbolic discounting” *Quarterly Journal of Economics*, 112 (2), 443-478.

Loewenstein G. and D. Prelec (1992) “Anomalies in intertemporal choice: evidence and an interpretation” *Quarterly Journal of Economics* 107 (2), 573-597.

Mas-Colell A., M. Whinston and J. Green (1995), Microeconomic Theory. New York: Oxford Press.

Mazur J. E. (1987) “An adjustment procedure for studying delayed reinforcement” in *The Effect of Delay and Intervening Events on Reinforcement Value.* Michael L. Commons, James E. Mazur, John A. Nevin and Howard Rachlin, eds. Hillsdale, NJ: Erlbaum.

McClure S. M., Ericson K. M., Laibson D. I., Loewenstein G. and J. D. Cohen (2007) “Time discounting for primary rewards” *Journal of Neuroscience* 27 (21), 5796-5804.

Metcalfe J. and W. Mischel (1999) “A hot/cool-system analysis of delay of gratification: Dynamics of willpower” *Psychological Review* 106 (1) 3-19.

Mischel W. and E. B. Ebbesen (1970) “Attention in delay of gratification” *Journal of Personality and Social Psychology* 16 (2), 329-37.

Mischel W., Ebbesen E. B. and A. Raskoff Zeiss (1972) “Cognitive and attentional mechanisms in delay of gratification” *Journal of Personality and Social Psychology* 21 (2), 204-21.

Mogg K. B., Field B. P., and J. De Houwer (2003) “Eye movements to smoking-related pictures in smokers: Relationship between attentional biases and implicit and explicit measures of stimulus valence” *Addiction*, 98, 825-836.

O’Donoghue T. and M. Rabin (2015) “Present bias: Lessons learned and to be learned” *American Economic Review: Papers & Proceedings*, 105 (5), 273-279.

Petry N. M., Bickel W. K and M. Arnett (1998) ”Shortened time horizons and insensitivity to future consequences in heroin addicts“, *Addiction* 93 (5), 729-738.

Phelps E. S. and R. A. Pollak (1968) “On second-best national saving and game-equilibrium growth” *Review of Economic Studies*, 35 (2), 185-199.

Radu P. T., Yi R., Bickel W. K., Gross J. J. and S. M. McClure (2011) “A mechanism for reducing delay discounting by altering temporal attention” *Journal of the Experimental Analysis of Behavior*, 96 (3), 363-385.

Read D. (2001) “Is time-discounting hyperbolic or subadditive?” *Journal of Risk and Uncertainty*, 23 (1), 5-32.

Read D. and P. Roelofsma (2003) “Subadditive versus hyperbolic discounting: A comparison of choice and matching" *Organizational Behavior and Human Decision Processes* 91 (2), 140-53.

Samuelson P. (1937) "A Note on measurement of utility" *Review of Economic Studies* 4 (2), 155-61.

Scholten M. and D. Read (2006) "Discounting by intervals: A Generalized Model of Intertemporal Choice" *Management Science* 52 (9), 1424-36.

Stout G. F. 1932. A Manual of Psychology. London: W. B. Clive, University Tutorial Press Ltd.

Taatgen N. A., Hedderik von R. and J. Anderson (2007) "An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning" *Psychological Review* 114 (3), 577-598.

Thomas E. A. and T. I. Brown (1974) "Time perception and the filled-duration illusion" *Perception and Psychophysics*, 16, 449-458.

Thomas E. A. and W. B. Weaver (1975) "Cognitive processing and time perception" *Perception and Psychophysics*, 17, 363-367.

Townshend J. and T. Duka (2001) "Attentional bias associated with alcohol cues: Differences between heavy and occasional social drinkers" *Psychopharmacology*, 157, 67-74.

Vuchinich R. E. and C. A. Simpson (1998) "Hyperbolic temporal discounting in social drinkers and problem drinkers" *Experimental and Clinical Psychopharmacology* 6 (3) 292-305.

Zakay D. and R. A. Block (1997) "Temporal cognition" *Current Directions in Psychological Science*, 6, 12-16.

Zauberman G. B., K. Kim, S. Malkoc and J. R. Bettman (2009) "Discounting time and time discounting: Subjective time perception and intertemporal preferences" *Journal of Marketing Research* 46, 543-56.

Appendix Proofs of Propositions

Proof of Proposition 1 Using vNM Theorem 1 and A6 (time additivity), $(x,t;p) \succcurlyeq' (x,t;q)$ is equivalent to

$$\sum_{\Theta} P(x,t,\theta)u_{vNM}(x,t,\theta) \geq \sum_{\Theta} Q(x,t,\theta)u_{vNM}(x,t,\theta) \tag{10}$$

where $r \in T$ exists such that $\sum_{\Theta} P(x,t,\theta|\theta_r = 1) = \sum_{\Theta} Q(x,t,\theta|\theta_r = 1)$, or with short notations $p_r = q_r$. The only case in which the inequality is not affected by arbitrary variations of $p_r = q_r$ is $u_{vNM}(x,t,\theta) = \sum_{j=0}^{T} u_{add}(x,t,\theta_j)$, that is

$$\sum_{\Theta} P(x,t,\theta)u_{vNM}(x,t,\theta) = \sum_{\Theta} P(x,t,\theta)\sum_{j=0}^{T} u_{add}(x,t,\theta_j)$$
$$= \sum_{j=0}^{T}\left(\sum_{\Theta} P(x,t,\theta|\theta_j = 1)u_{add}(x,t,\theta_j = 1) + \sum_{\Theta} P(x,t,\theta|\theta_j = 0)u_{add}(x,t,\theta_j = 0)\right)$$
$$= \sum_{j=0}^{T}\Big(p_j u_{add}(x,t,\theta_j = 1) + (1-p_j)u_{add}(x,t,\theta_j = 0)\Big)$$

Moreover, A8 (temporal indifference) implies for all $x \in X$, $t, s \neq t \in T$ $u_{add}(x,t,\theta_s = 0) = 0$ and A7 (waiting aversion) $u_{add}(x,t,\theta_s = 1) < 0$ for all $s < t$ and $u_{add}(x,t,\theta_s = 1) = 0$ for all $s > t$. A4 (monotonicity) implies $u_{add}(x,t,\theta_t) > 0$ for all $\theta_t \in \{0,1\}$ and A9 (static indifference) $u_{add}(x,t,\theta_t = 0) = u_{add}(x,t,\theta_t = 1)$.

To repeat with shorter notations: if reminding, waiting costs are $u_{add}(x,t,\theta_j = 1) = u_{add}(x,t,j) < 0$, $j < t$ and discounted utility of consumption $u_{add}(x,t,\theta_t = 1) = u_{add}(x,t,t) > 0$. If not, $u_{add}(x,t,\theta_j = 0) = 0$ if $j \neq t$ and $u_{add}(x,t,\theta_t = 0) = u_{add}(x,t,t)$. The vNM Theorem 4.1 simplifies to $(x,t,p) \succcurlyeq' (y,s,q)$ if

$$u_{add}(x,t,t) + \sum_{j=0}^{t-1} p_j u_{add}(x,t,j) \geq u_{add}(y,t,t) + \sum_{j=0}^{t-1} q_j u_{add}(y,t,j)$$

Now, according to A5 (Thomsen separability), for all $x,y,z \in X$, $t,s,r \in T$ and

$p \in [0,1]^{\bar{t}+1}$,

$$\frac{u_{add}(x,t,t) + \sum_{j=0}^{t-1} p_j u_{add}(x,t,j)}{u_{add}(y,s,s) + \sum_{j=0}^{s-1} p_j u_{add}(y,s,j)} = \frac{u_{add}(z,t,t) + \sum_{j=0}^{t-1} p_j u_{add}(z,t,j)}{u_{add}(y,r,r) + \sum_{j=0}^{r-1} p_j u_{add}(y,r,j)} = 1$$

is equivalent to

$$u_{add}(x,r,r) + \sum_{j=0}^{r-1} p_j u_{add}(x,r,j) = u_{add}(z,s,s) + \sum_{j=0}^{s-1} p_j u_{add}(z,s,j)$$

iff $u_{add}(x,t,j) = u(x)\delta(t,j)$ for all $x \in X$ and $t \in T$. The vNM Theorem 4.1 simplifies further to $(x,t,p) \succcurlyeq' (y,s,q)$ if

$$\delta(t,t)u(x) + \sum_{j=0}^{t-1} p_j \delta(t,j)u(x) \geq \delta(s,s)u(y) + \sum_{j=0}^{s-1} q_j \delta(s,j)u(y)$$

$\delta(t,j)u(x)$, $j \neq t$ are date j waiting costs, proportional to deferred utility, whereas $\delta(t,t)u(x)$ is discounted utility of consumption. According to A8 (Temporal indifference), for all $t \in T - \{\bar{t}\}$, if $p_t = 0$

$$\delta(t,t) + \sum_{j=0}^{t-1} p_j \delta(t,j) = \delta(t+1,t+1) + \sum_{j=0}^{t-1} p_j \delta(t+1,j)$$

For $t = 0$, consuming x at date 0 or 1 are equivalent if $\delta(0,0) = \delta(1,1)$. Consuming x at date 1 or 2 are equivalent if $\delta(2,2) - \delta(1,1) + p_0\big(\delta(2,0) - \delta(1,0)\big) = 0$ for all $p_0 \in [0,1]$, hence $\delta(1,0) = \delta(2,0)$ and $\delta(1,1) = \delta(2,2)$. Likewise, indifference between dates $t = 2, ..., T-1$ and $t+1$, with $p_t = 0$, implies for all $\{p_0, p_1, ..., p_{t-1}\} \in [0,1]^t$:

$$\delta(t+1,t+1) - \delta(t,t) - p_0\big(\delta(t+1,0) - \delta(t,0)\big) - p_1\big(\delta(t+1,1) - \delta(t,1)\big)$$
$$- \ldots - p_{t-1}\big(\delta(t+1,t-1) - \delta(t,t-1)\big) = 0$$

Hence, $\forall t = 1, ..., T-1$, $\delta(t,t) = \delta(t+1,t+1)$, and $\delta(t,s) = \delta(t+1,s)$, $\forall s = 0, ..., t-1$. Blocking the reminding date $s = 0, ..., T-1$ and varying the consumption date $t = 1, ..., T$ gives $\delta(s+1, s) = \delta(s+2, s) = ... = \delta(T, s)$. Normalizing $\delta(t, t) = 1$ for all $t \in T$ and simplifying notations: $\delta(t, s) = -\delta_s$ for all $s, t = 1, ..., T$, $s \neq t$, it follows that $(x, t, p) \succcurlyeq' (y, s, q)$ if

$$u(x) - u(x) \sum_{j=0}^{t-1} p_j \delta_j \geq u(y) - u(y) \sum_{j=0}^{s-1} q_j \delta_j$$

Next, according to A4 (Monotonicity), for all $x, y \in X$, $y > x$, $t \in T$ and $p \in [0, 1]^{\bar{t}+1}$

$$\big(u(y) - u(x)\big)\Big(1 - \sum_{j=0}^{t-1} p_j \delta_j\Big) > 0$$

which is satisfied if, setting all probabilities to 1, $1 - \sum_{j=0}^{t-1} \delta_j > 0$. Moreover, since all $\delta_{t-1} > 0$ (A7), we have $1 > 1 - \delta_0 > ... > 1 - \delta_0 - ... - \delta_{T-1} > 0$. □

Proof of Proposition 2 $D(t) = 1 - p_0\delta_0 - p\delta_1 - ... - p\delta_{t-1}$ (Proposition 1). A11 implies $p_j\delta_j + p_{j+1}\delta_{j+1} = (p_j + \Delta)\delta_j + (p_{j+1} - \dfrac{\Delta}{\beta})\delta_{j+1}$, for all $j < t - 2$ and $\Delta \in (0, min(1 - p_j, \beta p_{j+1}))$, hence $\delta_{j+1} = \beta\delta_j$, or $\delta_{j+1} = \beta^{j+1}\delta_0$. A10 implies $\beta \in (0, 1)$. Last, $D(t) > 0$ for all $t \in T$ and $p \in [0, 1]^{\bar{t}+1}$ (A4). $D(t)$ is minimal for $t = \bar{t}$ and $p_j = 1$ for all $j \in T$, hence $\delta_0 < \dfrac{1-\beta}{1-\beta^{\bar{t}}}$. □

Proof of Proposition 3 Applying Proposition 2 with all $p_j = \pi$ except $p_0 = 1$ (Assumption 2): $D(0) = 1$, $D(1) = 1 - \delta_0 = \beta$ and for $t > 1$

$$\begin{aligned} D(t) &= 1 - \delta_0 - \pi\beta\delta_0 - \pi\beta^2\delta_0 - ... - \pi\beta^{t-1}\delta_0 \\ &= 1 - (1-\pi)\delta_0 - \pi\delta_0(1 + \beta + \beta^2 + \beta^{t-1}) \\ &= 1 - (1-\pi)\delta_0 - \pi\delta_0 \frac{1-\beta^t}{1-\beta} \end{aligned}$$

With Assumption 1: $D(t) = 1 - (1-\pi)\delta_0 - \pi(1-\beta^t) = \pi\beta^t + (1-\pi)\beta$. □

Proof of Proposition 4 The indifference conditions $(x, 0) \sim (z, s)$, $(z, s) \sim (y, t)$ and $(x, 0) \sim (y', t)$ express as:

$$u(x) = (1 - p_0\delta_0 - p_1\delta_1 - ... - p_{s-1}\delta_{s-1})u(z)$$

$$(1 - p_0\delta_0 - p_1\delta_1 - ... - p_{s-1}\delta_{s-1})u(z) = (1 - p_0\delta_0 - p_1\delta_1 - ... - p_s\delta_s - ... - p_{t-1}\delta_{t-1})u(y)$$

$$u(x) = (1 - p_0\delta_0 - p_1\delta_1 - ... - p'_s\delta_s - ... - p_{t-1}\delta_{t-1})u(y')$$

The equalities simplify to $(1 - p_0\delta_0 - ... - p'_s\delta_s - ... - p_{t-1}\delta_{t-1})u(y') = (1 - p_0\delta_0 - ... - p_s\delta_s - ... - p_{t-1}\delta_{t-1})u(y)$. $y > y'$ if $p_s > p'_s$. □