An Investigation of Time Consistency for Subjective Discounted Utility

André Lapied
GREQAM
Aix-Marseille School of Economics

Olivier Renault
GREQAM
Aix-Marseille School of Economics

Abstract

A well-known common agreement in decision theory is that only exponential decision makers are time consistent i.e. with the mere passage of time, future choices must not contradict the initial choice. Building on this result, a large range of works has studied time inconsistency as a direct application of hyperbolic discounting. These articles share the common objective time assumption under which decision makers have a perfect perception of future periods. This paper firstly highlights that, when no further condition than separability is mentioned, any discount mechanism is compatible with time consistency. Then, we investigate time consistency assuming that individual time perception may be submitted to time distortion. In particular, an axiomatic discounting model called Subjective Discounted Utility (SDU) is provided to illustrate how hyperbolic decision makers may be time consistent.

KEYWORDS: Discounting, Time Distance, Time Consistency, Separable Preferences

JEL classification: D90

Introduction

Most of intertemporal decision models have focused on discounting, namely a psychological mechanism from which the decision maker translates her future instantaneous utility into a present instantaneous utility scale. Discounting has huge economic implications as a whole: investment project, retirement strategy and environmental policy are salient examples that highlight the extensive ramifications of discounting to various fields of analysis. Economists have been firstly attracted by exponential discounting for its simplicity, since discount rates between two successive periods are assumed to be constant. Exponential discounting, associated with separability, is the central mainstay of the standard exponential discounted utility model (Samuelson (1937), Koopmans (1960)). But for a few decades, exponential discounting has been challenged by alternative behaviors called hyperbolic discounting. Empirical studies (Thaler (1981)) first revealed hyperbolic discounting whose key property is to catch decreasing impatience: hyperbolic agents exhibit steeper discounting for time intervals closer to the present. Frederick, Loewenstein and O'Donoghue (2002) pointed out the descriptive limits of exponential discounting and underlined the great diversity in elicited discount rate. Building upon these experimental evidences, theoretical works such as Loewenstein and Prelec (1992), Harvey (1995), Laibson (1997)[1] or Prelec (2004) have contributed to precisely define the meaning of decreasing impatience and hyperbolic discounting is now a standard concept in decision theory. Paradoxically, hyperbolic discounting did not succeed in replacing exponential discounting in most of theoretical models. The reason is directly related to the following question: if a decision maker (DM) has the possibility to revisit her previous choice, will she confirm the initial decision? If the answer is affirmative, then the DM is said to be time consistent in the sense that preferences reversal cannot occur with the mere passage of time. Conversely, if the answer is negative, the DM is said to be time inconsistent. Time inconsistency rises in situation of temptation or procrastination and traditional applications are undersaving (Laibson (1997)) or addiction (Becker (1988)). Prelec (2004) stressed that time inconsistency induces inefficient choices due to unintended reversal of preferences for naïve agents, i.e. blinded individuals who do not anticipate the possibility to reconsider later their initial choice. That is why time consistency is traditionally viewed as a normative condition and as a firmly rooted component of intertemporal rationality. Since Strotz's seminal article (1956), it is well known that under a specific assumption called algebraic distance, a decision maker is time consistent if and only if she is exponential. Strotz's theorem has had a deep influence in intertemporal theory so that this conditional result (i.e. *based on* the algebraic distance) has progressively shifted toward an absolute "dogma" (i.e. *dissociated from* the algebraic distance) and the algebraic distance assumption has been progressively forgotten. This abusive common consensus has been designated by Drouhin (2009) as a "conventional wisdom" and will be called in this paper the Intertemporal Folk Theorem (IFT). It is worth noticing that while radical changes have been introduced into intertemporal theory for a few decades – and especially the growing success of hyperbolic discounting –, the IFT has still remained a

[1] Laibson has developed a quasi hyperbolic function where there is only decreasing impatience for the first period and for all other period the decision maker satisfies constant impatience.

deeply rooted result, so that it is now a major cornerstone of intertemporal decision. The consensus conveyed by the IFT solves the previous paradox on discounting by imposing two restrictive interpretations. First the IFT confers to exponential discounting an absolute and exclusive relation with time consistency. Indeed, the IFT attributes to exponential behavior an unexpected normative agreement and, in spite of a poor descriptive power, exponential discounting keeps an unyielding comparative advantage on hyperbolic discounting. Second the IFT implies that hyperbolic discounting appears as the exclusive preferential cause to analyze misbehaviors related to time inconsistency. Given the great implications associated with Strotz's theorem, it is quite surprising that the algebraic distance gave rise to so slight interest in intertemporal decision. As a matter of fact, the algebraic distance is often substituted by a Stop Watch Time assumption equivalent to time preservation, namely, "the zero of time is resetting to the moment of decision". However we will show that the algebraic distance is a stronger assumption than Stop Watch time since the algebraic distance is equivalent to both objective time and time perception. Objective time is an implicit but decisive hypothesis shared by the great majority of intertemporal models (Phelps and Pollak (1968), Harvey (1994), Laibson (1997) or Harris and Laibson (2001)). Objective time means that the DM has a perfect perception of future time intervals. For instance, the anticipated time duration of the coming month must be equal to the anticipated time duration for a month postponed in a decade. Nevertheless, as first noticed by A.C. Pigou, individuals often have "defective telescopic" ability to estimate far time intervals. In others words, individual's time perception may be submitted to time distortion and near time intervals may for example seem larger than equal more delayed time intervals. Albrecht and Weber (1995) introduced time perception into a discounted utility model. They assumed that exponential discounting is associated to objective time perception and then shown that deviation from standard discounting is fully explained by time perception distortion. In this way, hyperbolic agents have a logarithmic time perception. Zauberman and *ali.* (2009) claimed for a larger audience dedicated to time perception in intertemporal choices and Lapied and Renault (2012) indicated that time perception is a mean to take into account for the great diversity in elicited discount rates. In this paper we plan to reconsider the Intertemporal Folk Theorem under the light of time perception. In particular, in reference to Strotz's seminal article, we introduce the concept of time distance as a dynamic version of time perception. Indeed, a time distance measures future time intervals depending on the observational decision period. Thus time distance both covers the definition of the time perception for the initial decision period, but also the transmission of time perception properties to the following decision period. For instance, the perception of future periods may be sensitive to the age of the DM, since time might seem running faster when the DM is getting older. The algebraic distance is then a special case of time distance for which no time distortion occurs. We build a model called Subjective Discounted Utility (SDU) in which the decision maker faces time in two distinct postures: time distance which associates to each objective future period a unique subjective future period and discounting that converses subjective future into present scale. By analogy to the Subjective Expected Utility (SEU) for risky choices, the Subjective Discounted Utility is a two-stage-decision model in which the DM discounts subjective periods. The main result associated to the SDU is that time consistency is reduced to the specific class of time distances that satisfies a triangular equality condition.

This paper is organized as follows. Section 2 provides notations and a general representation with separability of time preferences. Section 3 will show that, without any additional condition, time consistency is consistent with all discount functions. The central interest of section 4 is to introduce time distances into discounted utility representation to reinterpret the time consistency condition and, building on these results, section 5 develops the Subjective Discounted Utility to propose a simple rule which characterizes time consistency.

Section 2 The Setting

2.1 Notations

A decision maker is first choosing between temporal prospects, i.e. sequences of future outcomes spread off over time. The income repartition between consumption and saving for the next year, the monthly allocation of time resource between labor and leisure and the organization for the next holiday week of pleasant and unpleasant tasks are salient examples of temporal prospects. With the mere passage of time, the DM has the possibility to reconsider later her initial plan. We assume a consequentialist naïve DM namely she has only a concern about future periods and does not anticipate the future opportunity to reconsider her first choice. The core interest of the paper is to determine the condition under which future choices do not contradict initial decision. The period set $\mathcal{P}$ is the set of non negative real numbers (with 0 as the initial period) and the DM's time horizon shall be $\mathcal{P}_i$ the set of non negative real numbers higher or equal to $i \in \mathcal{P}$. For any decision period i and given an index set $S = \{0,1,2, \dots, k, \dots\}$, the time sequence set is $\mathbb{T}_i = \{\tau : S \to \mathcal{P}_i$ with $\forall k \in S, i \leq \tau_k < \tau_{k+1}\}$ and the outcome sequence set is $\mathbb{X} = \{c : S \to \mathcal{C}\}$ where $\mathcal{C}$ is a compact topological space[2]. The decision set for period i is Δ_i then $\Delta_i = \mathbb{X} \times \mathbb{T}_i$ where a choice alternative is a sequence of dated rewards as follows $(x; t) = \{(c_0, \tau_0), (c_1, \tau_1), (c_2, \tau_2), \dots\}$. The decision set Δ_i, endowed with the product topology, is connected and separable. A binary relation $\succsim_i$ is defined on $\Delta_i \times \Delta_i$ for each decision period. As usual, $\succ_i$ and $\sim_i$ are defined to be respectively the asymmetric and symmetric parts of $\succsim_i$. For $z \in \mathcal{C}$, $(c; \tau) \in \Delta_i$, the notation $(c_k z; \tau) = \{(c_0, \tau_0), \dots, (c_{k-1}, \tau_{k-1}), (z, \tau_k), (c_{k+1}, \tau_{k+1}) \dots\}$ means that reward c_k for period τ_k of the outcome sequence $(c; \tau)$ (with $\tau_k \in \tau$) is replaced with the reward z at the same period. Similarly the notation $(c_{k,l} w, z; \tau)$ means that rewards c_k and c_l for periods τ_k and τ_l of the outcome sequence $(c; \tau)$ (with $\tau_k, \tau_l \in \tau$) are respectively replaced with rewards w and z at the same periods and so on. Lastly, a structural condition on time preference as a whole restricts our attention to the subset $\Delta_i{}^*$ which gathers all the choice alternatives $(c; \tau)$ verifying the following property: for any $\varepsilon > 0$, there exists an integer N such that for all $n \geq N$

$$\{(c_0 + \varepsilon, \tau_0), (c_1, \tau_1), \dots, (c_n, \tau_n), (0, \tau_{n+1}), (0, \tau_{n+2}), \dots\} \succsim_i (c; \tau)$$

[2] As a consequence the set of possible outcomes is assumed the same for all periods.

$$(c;\tau) \succsim_i \{(c_0-\varepsilon,\tau_0),(c_1,\tau_1),\dots,(c_n,\tau_n),(0,\tau_{n+1}),(0,\tau_{n+2}),\dots\}$$

As stressed by Harvey (1995), "the subset $[\Delta_i{}^*]$ is the larger set on which an additive function converges and is a value function for the preference relation".

2.2 Representation of Time Preferences with Separability

We propose a representation theorem for time preferences mainly based on a separability axiom.

A1 (*Regularity*)

A regular binary relation on $\Delta_i{}^*$ is a continuous monotonic myopic weak order[3] and has neutral element c^* in $\mathbb{X}$ such that $\forall \tau,\tau' \in \mathbb{T}_i$ and for all subset $A \subseteq S$, $(c^*{}_A z;\tau) \sim_i (c^*{}_A z,\tau')$ (with $z : A \to \mathcal{C}$).

A2 (*Separability*)

While regularity has to be seen as minimal condition for intertemporal rationality, separability is a strong assumption. Indeed, separability conveys the ideas that any dated reward can be estimated independently from each other (sequential independence) and that time estimation for dated reward is independent of what the reward is (magnitude independence).That is why separability axiom shall be divided into two distinct parts.

A2a (*First Order Separability*)

For all $w,x,y,z \in \mathcal{C}$, $c,c' \in \mathbb{X}$, $\tau \in \mathbb{T}_i$ and $i \in \mathcal{P}$

$$(c_{k,k+1}x,y;\tau) \succsim_i (c_{k,k+1}w,z;\tau) \Rightarrow (c'_{k,k+1}x,y;\tau) \succsim_i (c'_{k,k+1}w,z;\tau)$$

First Order Separability (FOS) insures the absence of complementarities across successive periods[4]. Associated with regularity, FOS implies that the order between outcome sequences $(c^*{}_s x;\tau)$ can be entirely deduced from the order between dated rewards (x,s). In the remaining analysis, we will use a slight abuse of notation for preferences defined on $\mathcal{C} \times \mathcal{P}_i$ writing $(x,s) \succsim_i (y,t)$ and keeping in mind that $(x,s) \succsim_i (y,t)$ means $(c^*{}_s x;\tau) \succsim_i (c^*{}_t y;\tau)$. The proof of next theorem 2.1 will show that A1 implies the existence of an intertemporal utility function U_i on Δ_i and then FOS allows the restriction of the analysis to instantaneous utility functions $v_i(x,t)$ defined on $\mathcal{C} \times \mathcal{P}_i$.

A2b (*Second Order Separability*)

For all $x,y,z \in \mathcal{C}$, $\tau \in \mathbb{T}_i$ and $i,j \in \mathcal{P}$ $(i<j)$,

$$\text{If } \begin{matrix}(x,s) \sim_i (y,t)\\ (x,s') \sim_j (y,t')\end{matrix} \text{ then } (c^*{}_s z;\tau) \sim_i (c^*{}_{s,t}x,y;\tau) \Rightarrow (c^*{}_{s'}z;\tau) \sim_j (c^*{}_{s',t'}x,y;\tau)$$

[3] The precise definition of a regular preference will be provided in Appendices.

[4] For models that integrate some complementarity between successive periods, see Choquet Utility Models as in Gilboa (1989), de Waegenaere and Wakker (2001) or Chateauneuf and Rébillé (2004).

Second Order Separability (SOS) is a traditional separability assumption as the Thomsen condition used by Fishburn and Rubinstein (1982). However SOS says a little more since it implies that temporal aspects, i.e. both period t and decision period i, of dated rewards may be isolated from outcome's estimation. We do not claim for a high descriptive degree associated with separability. Rather, assuming a very standard approach confers to our next results a stronger significance. We turn now to the third axiom retained for our representation theorem.

A3 (*Impatience*)

$\forall x, y > 0, \forall s, t \in \mathcal{P}_i\ (j < t)$ and $\forall i, j \in \mathcal{P}\ (i < j)$

$$(*) \qquad (0, s) \sim_i (0, t)$$

$$(**) \qquad s < t \Longrightarrow (x, s) \succ_i (x, t)$$

$$(***) \qquad (x, i) \sim_i (y, t) \Longrightarrow (y, t) \succ_j (x, j)$$

It is now usual to consider impatience as a core element of the intertemporal rationality. Impatience says that the receipt of the zero reward is matter of indifference (*), that the DM prefers to receive desired outcomes earlier than later (**) and that the mere passage of time increases the desirability of higher postponed outcomes (***).

Definition 1 (Separable Intertemporal Utility Functions)

A separable intertemporal utility function is a function that exhibits additivity between dated rewards and magnitude independence for any dated reward. Thus, for all $(c; \tau) \in \Delta_i^*$ and $i \in \mathcal{P}$,

$$U_i(c; \tau) = \sum_{k=0}^{+\infty} v_i(c_k, \tau_k) \qquad (1)$$

$$v_i(c_k, \tau_k) = m(i; \tau_k) u(c_k) \quad (2)$$

The next theorem proposes a special functional representation based on separable intertemporal utility function as in definition 1 for time preferences which satisfy the previous set of axioms.

Theorem 2.1 (*Representation Theorem with Separability*)

Let $\succsim_i$ be a preference relation on Δ_i^* and assume that $(c; \tau), (c'; \tau') \in \Delta_i^*$. Then the following two statements are equivalent:

(i) *The binary relation $\succsim_i$ verify regularity, separability and impatience*
(ii) *The separable intertemporal utility function U_i is a representative function for $\succsim_i$*

$$(c; \tau) \succsim_i (c'; \tau') \Leftrightarrow U_i(c; \tau) \geq U_i(c'; \tau')$$

$$U_i(c; \tau) = \sum_{k=0}^{+\infty} m(i; \tau_k) u(c_k) \qquad (3)$$

The instantaneous utility function u[5] is unique up to a positive linear transformation with $u(0) = 0$. Furthermore, decision weights satisfy $m : \mathcal{P} \times \mathcal{P}_i \to]0;1]$, $m(i;i) = 1$, $m(i;t_k) > 0$, m is increasing with i and decreasing with t_k.

Section 3 Discounting and Time Consistency

3.1 Dynamic Discount Functions

For a few decades, literature dedicated to discounting has been producing a substantial flow of studies. Empirical studies reveal a sizeable diversity in elicited discount rates and the most frequent observed feature is decreasing impatience (Frederick, Loewenstein and O'Donoghue (2002)) for hyperbolic agents. Theoretical works have integrated decreasing impatience in axiomatic models (Loewenstein and Prelec (1992)) so that hyperbolic discounting is now very standard in intertemporal choices. Nevertheless, in spite of a poor descriptive power, exponential discounting (or constant impatience) paradoxically remains the privileged way for economic applications dealing with optimal behavior.

Definition 2 *(Discount Functions)*

A discount function $\mu : \mathcal{P} \to]0,1]$ is a continuous function which simultaneously satisfies the three following properties:

(d1) $\mu(0) = 1$ (Normalization)

(d2) $\mu(t_k) > 0$ (Positive discounting)

(d3) μ is (weakly) decreasing with t_k (Impatience)

Property (d1) says that there is no discount for the present period and (d2) implies that the DM has minimal concern for all future periods, no matter how far from the present are. Lastly (d3) exhibits impatience.

Definition 3 *(Discounted Utility)*

We call Discounted Utility (DU) a separable intertemporal utility function whose decision weights are discount functions as presented in definition 2.

Observation 3.1

The intertemporal function (3) is a Discounted Utility (DU) since $\forall t \in \mathcal{P}_i$, decision weights are dynamic discount functions:

$$m(i;t) = \mu_i(t-i) \quad (4)$$

[5] Observe that u is independent of time. The critical point is that the instantaneous utility must be independent from the decision period i. As noticed by Drouhin (2009), instantaneous utility independence from period τ_k is not a necessary condition and dropping this assumption does not perturb our next results.

The great heterogeneity in elicited discount rates can be expressed by the following axiom. Each discount mechanism is uniquely associated to an discount premium θ.

A4 (*Time Discounting*)

$\forall s, t \in \mathcal{P}, \exists \theta = \theta(t, k): \mathcal{P} \times \mathcal{P} \to \mathcal{P}$ such that:

$$(x, s) \sim_0 (y, t) \Longrightarrow (x, s + k) \sim_0 (y, t + k + \theta)$$

For $s = 0$, $(x, 0) \sim_0 (y, t) \Longrightarrow (x, k) \sim_0 (y, t + k + \theta)$ where $(x, 0)$ is the present equivalent to (y, t). The standard model is the Exponential Discounted Utility (EDU) with $\theta = 0$ and $\mu_0(t) = \lambda^t$ and expresses constant impatience. EDU has been challenged in descriptive and prescriptive studies by Hyperbolic Discounted Utility (HDU) with $\theta = tk$ and $\mu_0(t) = 1/(1 + t)$. Hyperbolic agents exhibit steeper discounting for time intervals closer to the present. To explain the previous paradox, we need to introduce the normative condition of time consistency.

3.2 Time consistent choices

There is an old tradition in intertemporal choice for associating hyperbolic discounting to time inconsistency (Thaler (1981), Laibson (1997)). Time inconsistency rises when the mere passage of time induces preferences reversal. For instance, Akerlof (1991) studied the case of procrastination for which an initial "planner" chooses to accomplish an unpleasant task in remote future (say one month), but future "doer" reveals preference for delaying the unpleasant task (say one additional month). Therefore, future doer continually contradicts the initial plan and procrastination leads to the perpetual delaying of unpleasant tasks. Hyperbolic discounting is generally seen as a preferential source of procrastination since short term preference exhibits higher discount rate than long term preference. On the contrary, time consistency guaranties the absence of preferences reversal across decision periods. In other words, with the mere passage of time, the future "doer" does not contradict the choice of the initial "planner". Prelec (2004) noticed that time inconsistency induces inefficient choices due to unintended reversal of preferences for naïve agents. As a consequence, time consistency appears as a sound normative condition for rational choices.

A5 (*Time Consistency*)

$\forall x, y \in \mathcal{C}, \forall i, j \in \mathcal{P}\ (i < j), \forall s, t \in \mathcal{P}_j$

$$(x, s) \succsim_i (y, t) \Leftrightarrow (x, s) \succsim_j (y, t)$$

Proposition 3.2 (*Time Consistent Decision Weights*)

A decision maker is time consistent if and only if, for all decision period i, decision weights as in theorem 2.1 are weighted discount functions such that $\forall i \in \mathcal{P}, \forall t \in \mathcal{P}_i$

$$m(i;t) = \frac{\mu_0(t)}{\mu_0(i)} \qquad (5)$$

Proposition 3.2 merely shows that without any additional condition, time consistency is compatible with any discount mechanism. For instance, assume $\mu_0(t) = \lambda^{t^\alpha}$ $(\alpha > 0)$ then corresponding time consistent decision weights will be $m(i;t) = \lambda^{t^\alpha - i^\alpha}$. Observe that the previous discount function covers the different cases for impatience[6] because if $a < 1$ (resp. $a > 1, a = 1$) then decreasing (resp. increasing, constant) impatience holds.

Corollary 3.3

Let $\succsim_i$ *verify time consistency. Then for decision weights defined as in theorem 2.1, the two following statements are equivalent:*

$(*)$ $\forall s,t \in \mathcal{P}_i, s \leq t, m(i;t) = m(i;s) \times m(s;t)$

(**) $\forall (c;\tau) \in {\Delta_j}^*, \forall i,j \in \mathcal{P}$ $(i < j), \exists \phi: \mathcal{P} \times \mathcal{P}_i \rightarrow [1;+\infty[$ such that $U_j(c;\tau) = \phi(i;j) \times U_i(c;\tau)$ with $\phi(i;j) = m(i;j)^{-1}$

However, as seen before, contrary to the message conveyed by proposition 3.2, it is usual to consider that only exponential decision makers are time consistent. The restrictive additional condition allowing such an exclusive relation is called Stop Watch Time and resets the zero on the discount function when the next decision arrives.

Definition 4 (*Stop Watch Time*)

$$\forall i \in \mathcal{P}, \{m(i;t)\}_{t \in \mathcal{P}_i} = \{m(0;t-i)\}_{t \in \mathcal{P}_i}$$

Stop Watch Time is a restrictive condition since it imposes that each DM has the same decision map for all decision periods. In other words, time preferences are age independent since the older DM conserves the same structure of time preferences than young DM's.

Observation 3.4 (*Exponential Consistent DM*)

Let $\succsim_i$ *verify A1-A3. Then if SWT holds, the decision maker is time consistent if and only if EDU holds.*

As a consequence, observation 3.4 solves the above paradox since, assuming SWT, time consistency confers unexpected comparative advantage to exponential discounting. Alternatively, hyperbolic discounting is then mobilized to analyze misbehaviors such as undersaving, addiction or self confidence. For example assume a hyperbolic agent which satisfies SWT then theorem 2.1 gives the following decision weights $m(i;t) = 1/(1+t-i)$. Thus preferences reversal may hold since the DM both exhibits $({c^*}_{1,2} - 100{,}180;\tau) \succ_0 (c^*;\tau)$ and $(c^*;\tau') \succ_1 ({c^*}_{1,2} - 100{,}180;\tau')$ with $\tau' = \tau - \{\tau_0\}$.

[6] Section 5 develops this point by introducing the Prelec's criterion which is a measure for decreasing impatience degree.

Section 4 Time Distances

Our departure point for this section is an observation by Drouhin (2009). Consider the following decision weight $m(i;t) = (1+i)/(1+t)$ then by proposition 3.2, m satisfies time consistency with $\mu_0(t) = 1/(1+t)$ and hyperbolic decision makers may be time consistent.

Like SWT for EDU, the objective of this section is to determine a specific condition under which HDU is the only preference pattern compatible with time consistency. More precisely, this section will provide a general condition defining the exclusive association between time consistency and a specific DU model.

The first economist dealing with time consistency for naïve agents was Strotz (1956). The key element of his demonstration is the algebraic distance which imposes that discount weights only depends on the algebraic difference between discounted period and decision period ($m(i;t) = \mu(t-i)$). Strotz's famous theorem states that under algebraic distance, only exponential decision makers are time consistent. Undoubtedly, there is a close connection between Strotz's theorem and Observation 3.4 so that intertemporal theory traditionally considers the two assumptions as equivalent. However, as we will see, the algebraic distance is a stronger assumption and contains SWT. Such an abusive equivalence is called in this paper "Intertemporal Folk Theorem". We argue that this abusive assimilation is due to the fact that like SWT, algebraic distance in Strotz's methodology is a pure formal assumption without any behavioral perspective. In order to clearly dissociate algebraic distance from SWT and to break with the Intertemporal Folk Theorem, we revive Strotz's seminal article by including the algebraic distance in a larger concept named time distance.

Time distance is a dynamic version of time perception. Time distance measures at decision period i the gap between decision period i and future period t. Therefore if $\rho(t)$ denotes the DM's time perception, then $d(i;t) = \rho_i(t-i)$ defines her time distance. Intertemporal theory allows a growing interest to time perception (Albrecht and Weber (1995), Zauberman and ali. (2009)). The introduction of time perception into standard discounting fulfills the wish of giving a sound behavioral basis to recent changes observed in time discounting and of explaining the great diversity in elicited discount functions. In this way, time distances capture two behavioral schemes. The estimation of one year from now may have sensibly different treatment than one year in one decade ($\rho(1) \neq \rho(11) - \rho(10)$). In addition the young DM's consideration for one year in one decade may be sensibly different from the same consideration brought by older DM ($\rho_i(11) - \rho_i(10) \neq \rho_j(11) - \rho_j(10)$ $(i < j)$). The first example suggests that for a given observational time point, our perception of future duration may be *sensitive to the distance from the observational time point* and for instance the DM is more or less unable to mentally represent far future. The second example highlights that our perception of future time intervals may be *sensitive to the observational time point* and consideration for far future depend on the age of the DM.

In spite of the great heterogeneity allowed by time distortion, time distance functions share a limited set of properties.

Definition 5 *(Time Distance Functions)*

A time distance function $d : \mathcal{P} \times \mathcal{P}_i \to \mathcal{P}$ is a continuous function which simultaneously satisfies the three following properties:

(p1) $d(i;t) = \rho_i(t-i) \geq 0$ *(Dynamic Time Perception)*

(p2) d *is decreasing with* i *and increasing with* t

(p3) $d(i;i) = 0$ *(Normalization for present period)*

Property (p1) says that time distance is a subjective period and subjectivity may depend on decision period. Property (p2) insures that time distance decreases (resp. increases) when the DM is moving closer to (resp. away from) the discounted period. Normalization implies that the present is not submitted to time distortion. Dynamic discount weights ($\mu_i(t-i)$), as in Observation 3.1, cannot be easily deduced from static discount weights ($\mu_0(t)$) yet. Indeed, observe that decision weights are discount functions of *period gap* and not only of *period*. Thus, period gap may be dependent from the decision period. The next observation will clarify this point.

Observation 4.1

Decision Weights of Theorem 2.1 are discount functions of time distance, that is, $\forall i \in \mathcal{P}, \forall t \in \mathcal{P}_i$,

$$m(i;t) = \mu\big(d(i;t)\big) \quad (6)$$

Observation 4.1 has many implications. It first says that dynamic intertemporal decision weights condense two distinct temporal postures: discounting and time perception. However, by doing this, no clear relation is established between discounting and time distance. Furthermore Observation 4.1 implies that time distance integrally subsumes the dynamical aspect of the model. As a consequence, time distance may be summarized by the combination of two complementary conditions: an initial condition which defines time perception for initial decision period and a transmission condition which determines how time perception is sensible to decision period.

Initial condition:

$$d(0;t) = \rho_0(t) \equiv \rho(t) \quad (7)$$

Transmission condition:

$$d(i;t) = \rho(t-i) + \varphi(i;t) \quad (8)$$

For initial condition, we focus on convexity or concavity properties of time perception. Salient cases are Objective Time ($\rho(t) = t$) or Logarithmic Time Perception ($\rho(t) = \ln(1+t)$). For transmission condition, if $\varphi(i;t) = 0$ then the time distance satisfies Time Preservation ($\rho_i(t-i) = \rho(t-i)$). For $\varphi(i;t) > 0$, we will say that the DM exhibits Time Compression and if $\varphi(i;t) < 0$ that the DM exhibits Time Extension.

Proposition 4.2

All time distance function is uniquely defined by a combination of an initial condition and a transmission condition.

For instance, the algebraic distance $d(i;t) = t - i$ and the quasi algebraic distance $d(i;t) = (t-i)/(1+i)$ share the Objective Time property but have different transmission conditions. The algebraic distance and the logarithmic time distance $d(i;t) = \ln(1+t-i)$ share the time preservation condition but reflect different time perceptions. Corollary 4.4 summarizes these observations. But in order to prove it, we need to introduce, as for discounting, a premium τ which will characterize all time distances. This is precisely the objective of the next axiom.

A6 (*Gap Estimation*)

$\forall i,k \in \mathcal{P}, \forall s,t \in \mathcal{P}_i, \exists \tau = \tau(i,k,t): \mathcal{P} \times \mathcal{P} \times \mathcal{P}_i \to \mathcal{P}$ such that:

$$(x,s) \sim_i (y,t) \Rightarrow (x, s+k) \sim_{i+k} (y, t+k+\tau)$$

A6 is a quite intuitive axiom. For $s = i$, $(x,i) \sim_i (y,t) \Rightarrow (x, i+k) \sim_{i+k} (y, t+k+\tau)$, the premium τ indicates how the coming time interval perception between periods i and t may be affected by the mere passage of time. Obviously if $\tau = 0$ then time preservation holds. In return, if $\tau < 0$ (resp. $\tau > 0$) then the DM exhibits time compression (resp. extension) and time seems running faster (resp. more slowly) with age increasing. The next proposition will show that A6 is actually imposing a transmission condition for time distance, especially by clarifying the intimate relationship between τ and φ.

Proposition 4.3

For a given time perception, the previous axiom fully determines the transmission condition associated to the time distance. Moreover, for all $i \in \mathcal{P}$ *and* $t \in \mathcal{P}_i$, $\varphi(i; i+t) = d(i; i+t) - d\big(i; i+t+\tau(0,i,t)\big)$.

Proposition 4.3 clearly shows that there exists a negative relation between τ and φ since when τ increases, (p2) from definition 5 implies that d also increases and φ decreases. Notice that the magnitude of the "elasticity of φ" to a change in τ depends on the nature of time perception. Under the light of proposition 4.3, we are now able to complete the message associated to proposition 4.2 by introducing the following corollary.

Corollary 4.4

$(*)$ *Algebraic Distance (AD) is the unique time distance satisfying objective time and time preservation*

$(**)$ *Quasi Algebraic Distance (QAD) is the unique time distance satisfying objective time and time compression with premium* $\tau = k \times (t - i)/(1 + i)$

$(***)$ *Logarithmic Time Distance (LTD) is the unique time distance satisfying logarithmic time perception and time preservation*

The previous section highlighted the restrictive implication of SWT. Observation 3.4 is often assimilated to Strotz's theorem so that SWT and the algebraic distance are traditionally viewed as equivalent assumptions. However this assimilation is abusive since the algebraic distance is a stronger assumption.

Observation 4.5

For separable intertemporal utility functions as presented in definition 1, Time preservation is equivalent to SWT.

Actually, SWT is only one part of the algebraic distance. We now turn out the main theorem of this section and explain why this abusive association is frequently observed in intertemporal decision theory.

Theorem 4.6

For each given discount mechanism μ_0*, there is a set of time distances such that the decision maker is time consistent.*

$$d(i;t) = \mu^{-1}\left(\frac{\mu_0(t)}{\mu_0(i)}\right) \quad (9)$$

The main theorem is the largest result for time consistency based on discounting and time distance. The counterpart of this generality is a relative indetermination since a single decision weight m as in (6) can be obtained from different combinations of discount function and time distance function. For example, $m(i;t) = (1 + i)/(1 + t)$ may formally result from the combination between exponential form for μ and the following time distance $d(i;t) = r^{-1}[\ln(1 + t) - \ln(1 + i)]$ or from the combination between hyperbolic form for μ and quasi hyperbolic distance $d(i;t) = (t - i)/(1 + i)$. The reason for that indetermination is the absence of a behavioral correspondence between the two temporal postures. Indeed, observe that in this approach, the function μ may be independently chosen from the time distance. In the next section, the time distance fully determines the nature of discounting. Actually, the expression (9) associates to each discount mechanism μ_0 a specific class of time distances such that the DM is time consistent. It is worth noticing that theorem 4.6 may be strengthened by imposing only one additional core property of the algebraic distance.

Corollary 4.7

If (TP) is retained for (9) then only exponential DM are time consistent $\mu_0(t) = \lambda^t$

If (OT) is retained for (9) then $\mu_0(t) = \mu(t)$*, and a DM is time consistent provided that* $d(i;t) = \mu^{-1}(\mu(t)/\mu(i))$

First part of corollary shows that the Intertemporal Folk Theorem is a very restrictive version of a larger condition associated to time consistency. As a consequence, when time preservation (or SWT) is dropped, EDU loses the exclusive association with time consistency. Second part of corollary shows that, in a standard approach (OT), all decision makers can be time consistent and each discount mechanism is linked with a unique transmission condition. For instance, hyperbolic agents are time consistent if and only if they have a quasi algebraic distance. However, the second part of corollary 4.7 still remains too restrictive since the diversity of time perceptions is denied. In other words, Objective Time prevents us from building a sound behavioral structure for time consistent decision makers. The main theorem and its corollary suggest that a serious track to generalize Strotz's Theorem is relaxing the objective time assumption. To achieve this goal, the introduction of time perception into intertemporal models appears as an appealing way to consider subjective time.

Section 5 Subjective Discounted Utility

Theorem 4.6 is the largest result dealing with time consistency and based on time distance and discounting. However, it is a pure formal result with poor behavioral interpretation. In this section, we present a model in which discounting is fully explained by the nature of time distance. In other words, decision weights can be expressed in the following sense: for $\lambda = e^{-r}$ $(r > 0), \forall t \in \mathcal{P}_i, i \in \mathcal{P}$

$$m(i;t) = \lambda^{d(i;t)} \quad (10)$$

The Subjective Discounted Utility (SDU) was introduced by Lapied and Renault (2012) for a unique decision period model. The main interest of such a model is that each discount function is associated to a unique time perception. In other words, the SDU approach confers to individual discounting a clear behavioral interpretation. For instance, a pioneer article by Albrecht and Weber (1995) demonstrates that, as the famous "law" of diminishing marginal utility, hyperbolic agents have a logarithmic time perception, i.e. the marginal perceived time is a decreasing function of time. The SDU is based on a subjective discounting axiom. We will show that this axiom is necessary to get the desired form (10) for decision weights defined as in Observation 3.1.

A7 (*Subjective Discounting*)

$\forall s, t \in \mathcal{P}, \exists \rho: \mathcal{P} \to \mathcal{P}$ such that:

$$(x, s) \sim_0 (y, t) \Rightarrow \left(x, \rho^{-1}(\rho(s) + \rho(k))\right) \sim_0 \left(y, \rho^{-1}(\rho(t) + \rho(k))\right)$$

The SD axiom implies a special representation for decision weights at initial period $m(0;t) = \lambda^{\rho(t)}$ (Lapied and Renault (2012)). Furthermore, if $s = 0$ then A7 becomes $(x, 0) \sim_0 (y, t) \Rightarrow (x, k) \sim_0 \left(y, \rho^{-1}(\rho(t) + \rho(k))\right)$. The meaning of A7 is quite easy to perceive. First, all discount axioms like A4 involve objective time in both members of the proposition. Conditions in definition 5 ensure that ρ^{-1} does exist and if ρ converts all real periods in subjective dates then ρ^{-1} produces the reciprocal transformation namely each subjective date is associated to a unique objective period. Furthermore the expression $\rho(t)+\rho(k)$ can be interpreted in a traditional way considering $\rho(k)$ as an additional subjective delay. To illustrate, suppose that a 90 immediate reward is the same than obtain 100 in one period (say one year) so $(90{,}0) \sim_0 (100{,}1)$. Then A7 implies for a two period decay that $(90{,}2) \sim_0 \left(100, \rho^{-1}(\rho(1)+\rho(2))\right)$. Of course, it is not necessary that $\rho^{-1}(\rho(1)+\rho(2)) = 3$. If the case, then $\rho(t)=t$, i.e. time perception is reduced to objective time[7] and A7 changes for stationarity: $(x,0) \sim_0 (y,t) \Rightarrow (x,k) \sim (y,t+k)$. The SDU Model reduces to the EDU model with $\mu(t) = \lambda^t$. With the same example, if $\rho(t) = \ln(1+t)$ and logarithmic time perception induces $\rho^{-1}(\rho(1)+\rho(2)) = 5$. The DM exhibits decreasing impatience and A7 becomes the central axiom used by Harvey (1995), $(x,0) \sim (y,t) \Rightarrow (x,k) \sim \left(y, t+k(1+t)\right)$ with $\mu(t) = (1+t)^{-r}$.

By Observation 4.1, a direct consequence will be that the expression of dynamic decision weights changes for the general form $m(i;t) = \lambda^{d(i;t)}$. By analogy to the Subjective Expected Utility (SEU) for risky choices, the Subjective Discounted Utility is a two-stage-decision model in which the DM discounts subjective periods. In this way, the parameter λ can be meaningfully interpreted as a market discounting rule and potential individual deviations from this rule are fully explained by time distortions (time perception and/or time transmission) allowed by time distance functions.

Theorem 5.1 (*Subjective Discounted Utility*)

Assume (3) holds. Then the binary relation $\succsim_0$ satisfies A7 if and only if, for all $(c;\tau) \in {\Delta_i}^*$, $0<\lambda<1$, the intertemporal utility function is

$$SDU_i(c;\tau) = \sum_{k=0}^{+\infty} \lambda^{d(i;\tau_k)} u(c_k) \quad (11)$$

It is worth noticing that under (11), exponential decision makers are time consistent if and only if the algebraic distance ($d(i;t) = t-i$) holds. Nevertheless, by theorem 4.6, hyperbolic agents may also be time consistent if and only if time distance is $d(i;t) = r^{-1}\ln((1+t)/(1+i))$. Then the Strotz's Theorem is a special case of time consistency in the Subjective Discounted Utility model.

Corollary 5.2

In the SDU model, for all decision periods, the time distance integrally determines the nature of discounting for decision weights defined as in Observation 3.1

[7] Actually time perception needs to be linear, that is $\rho(t) = a \times t$ with $a>0$ but it is easily to show that in the SDU model, linear time perceptions may be reduced to the case of objective time.

By theorem 5.1 and corollary 5.2, the SDU insures that there exists a one-to-one correspondence between time distance and dynamic discount weight. As a consequence, discount properties are fully explained by behavioral features conveyed by time distance functions. That is why the subjective discounted utility appears as a simple means to find out a general condition for time consistency exclusively based on time distance properties. Looking back to the two previous time distances, what is the key property shared by both exponential and hyperbolic consistent agents?

We start from the Intertemporal Folk Theorem from which *under the algebraic distance*, only exponential decision makers can be time consistent. Intuitively, if one of the two conditions of the algebraic distance is weakened - say for instance objective time - then we will show that time preservation must be weakened in a *very precise way* such that the DM remains time consistent. We first characterize time consistency in the SDU model by a simple rule based on time distance properties and figure 1 plotted below will illustrate the intimate relation between time perception (initial condition) and the transmission condition. Then, we will reinforce this central interpretation by developing the meaning of the expression "*very precise way*" with the help of time premiums θ and τ respectively presented in section 3 and section 4 of the study.

Recall that time distances are singular distance functions with respect to mathematical distances defined in the usual sense. In particular, the natural ordering introduced by time elapsing imposes that time distances are not symmetric. However, the singularity of time distance may also be highlighted by the use of the traditional triangular inequality property $d(i;t) \leq d(i;s) + d(s;t)$ shared by all "spatial" distances. Indeed, due to potential distortion, time distances permit that the reverse triangular inequality such as $d(i;t) > d(i;s) + d(s;t)$. For instance, the following the quasi algebraic distance $d(i;t) = (t-i)/(1+i)$ implies $d(i;t) > d(i;s) + d(s;t)$.

Definition 5 (*Triangular Inequality for Time Distances*)

For all $s, t \in \mathcal{P}_i, i \in \mathcal{P}$,

$$d(i;t) \begin{pmatrix} > \\ = \\ < \end{pmatrix} d(i;s) + d(s;t) \quad (12)$$

The traditional part of the triangular inequality ($\leq$) may be interpreted as usual. For the reverse triangular inequality ($>$), the following intuitive interpretation holds. Consider an agent who has to return with delay, say two months, some report. The agent plans to work during the second month and considers that this time interval is sufficient to do the job. After one month elapsing, the coming month seems to be shorter so that the DM begins to doubt her ability to return the report in the right delay.

In most of the studies dealing with time consistency, objective time and time preservation are simultaneously retained and the Intertemporal Folk Theorem holds. It is easy to show that Objective Time implies $d(i;t) - d(i;s) = d(i;i+t-s)$ and that time preservation implies $d(s;s+k) = d(i;i+k)$(for d well defined). By associating the two previous expressions, the triangular equality holds $d(i;t) = d(i;s) + d(s;t)$. However, Objective Time and Time Preservation have a poor descriptive basis and observations tend to confirm two alternative time postures. On the one hand, the observation of defective telescopic abilities of individuals tends to claim for logarithmic time perceptions. On the other hand, empirical studies frequently reveal time compression, i.e. there is a larger propensity for older agents to reduce anticipated coming time intervals than for younger agents. Hence, logarithmic time perception insures that $d(i;t) - d(i;s) < d(i;i+t-s)$ and time compression that $d(s;s+k) < d(i;i+k)$. Therefore if we consider the case for $k = t - s$ then $d(i;t) - d(i;s) < d(i;i+k)$ and $d(s;t) < d(i;i+k)$ simultaneously result. Assume in addition that the previous inequalities reveal exactly the same intensity difference in the sense that $d(i;t) - d(i;s) = d(s;t)$ so the triangular equality is achieved again (the precise meaning of this assumption will be explained in proposition 5.5). Thus exponential and hyperbolic time consistent decision makers seem to exhibit the same condition. After reinterpreting the triangular inequality uniquely on the basis of time perception (Observation 5.3), we will show that for subjective discounted utilities, time consistency is fully subsumed by the triangular equality property (Proposition 5.4).

Observation 5.3

For all $t \in \mathcal{P}_i$, $i \in \mathcal{P}$, the Triangular Inequality can be rewritten as the comparison between the time distance and the difference between time perceptions for periods t and i:

$$d(i;t) \begin{pmatrix} < \\ = \\ > \end{pmatrix} \rho(t) - \rho(i) \quad (13)$$

Proposition 5.4

In the SDU model, decision makers are time consistent if and only if the triangular equality is verified

$$TC \Leftrightarrow d(i;t) = \rho(t) - \rho(i)$$

Proposition 5.4 confers to time consistent agents a common behavioral property in terms of time perception. Time consistency clearly appears as a very restrictive condition since it imposes that the perceived gap between two periods is independent from the observational point. For example, if a year in one decade seems now very short then, after the mere passage of a decade, the coming year has to appear as short as the previous estimation (and obviously, DM exhibits time compression).

Figure 1. The relation between the initial condition and the transmission condition in the SDU model

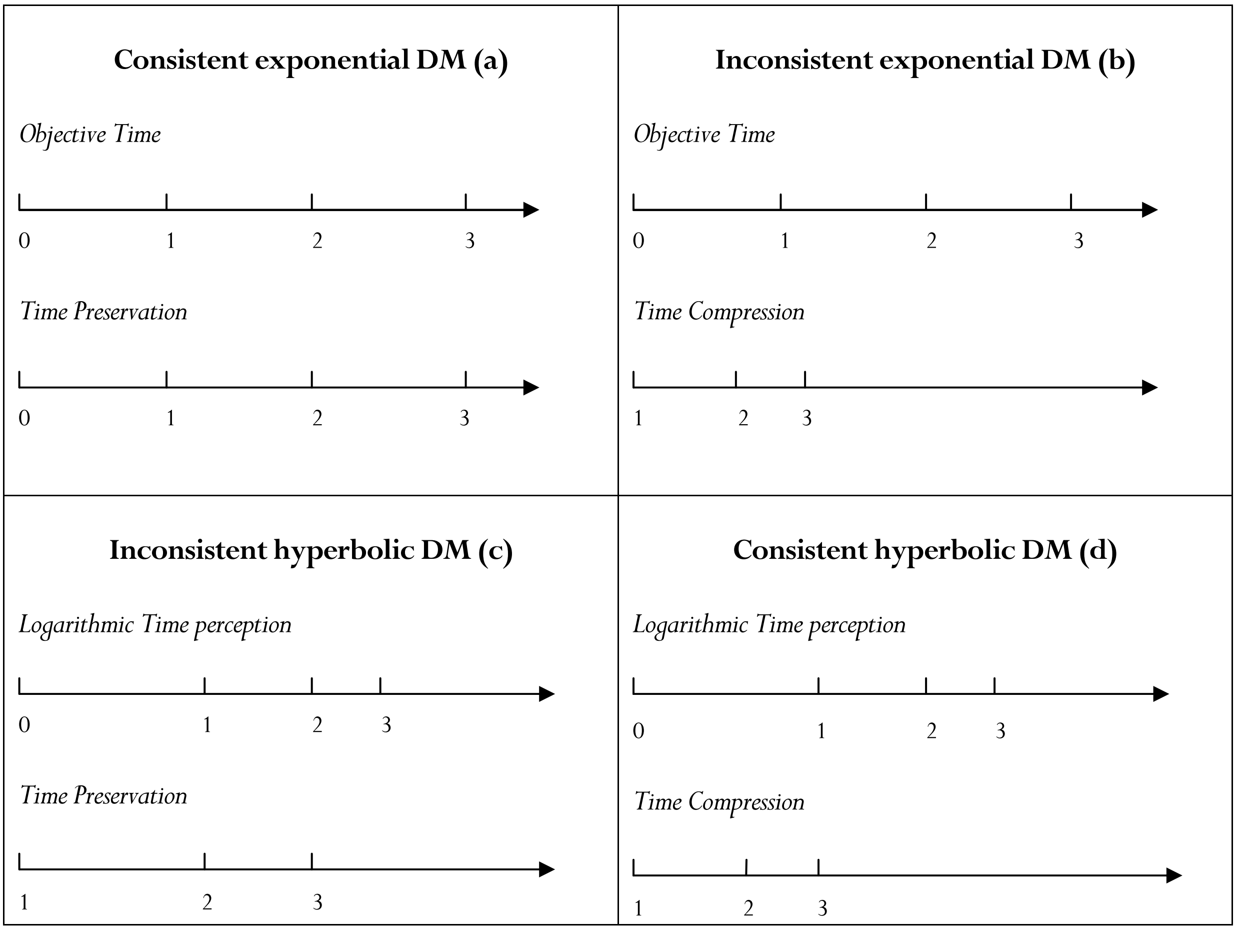

Figure 1 underline that, for any other initial condition other than Objective Time, changes in time perception must be counterbalanced by a symmetric effect in the transmission condition to preserve time consistency. Thus, by corollary 4.7, the introduction of the logarithmic time perception for consistent hyperbolic agents (switch from (a) toward (d)) is necessarily associated with time perception withdrawal. Moreover, remember that the logarithmic time perception implies that close future time intervals seem larger than equal more distant future time intervals. Therefore, in this case, the time consistency condition requires that the DM exhibits time compression so as to perfectly compensate the effect of logarithmic time perception (switch from (b) toward (d)). Obviously, another direct consequence of the Intertemporal Folk Theorem is that a DM who simultaneously exhibits Objective Time and other transmission condition than time preservation is time inconsistent (switch from (a) toward (b)).

At the end, we propose a simple condition to precisely define how the transmission condition counterbalances deviances from Objective Time to insure that non exponential decision makers are time consistent. First, Time discounting (A4) and Gap Estimation (A6) axioms from previous sections are for $s = 0$

A4 $(x, 0) \sim_0 (y, t) \Longrightarrow (x, k) \sim_0 (y, t + k + \theta)$ (with $\theta = \theta(t, k)$)

A6 $(x, 0) \sim_0 (y, t) \Longrightarrow (x, k) \sim_k (y, t + k + \tau)$ (with $\tau(0, k, t) \equiv \tau(k, t)$)

Immediately, the combination of A4 and A6 induces the next observation.

Observation 5.5

The DM is time consistent if and only if, for all $, t \in \mathcal{P}$ $(k \leq t)$, *the initial time premium* θ *equals transmission time premium* τ

$$TC \Leftrightarrow \theta(t, k) = \tau(k, t)$$

Indeed, Observation 5.5 will show that the time premium in A4 - that expresses, for instance, the additional delay needed to maintain indifference *for postponed rewards* - has to be equal to the time premium in A6 - that expresses, in this same case, the additional amount of time needed to preserve the indifference *for postponed decision*.

Notice that Observation 5.5 may be found independently from the SDU model. However, considering the subjective discounting axiom (A7) as a particular time discounting axiom for which $\theta(t, k) = \rho^{-1}(\rho(t) + \rho(k)) - (t + k)$, a key is introduced to understand the relationship between the standard interpretation of time consistency (based on time premiums which implicitly define discount rates) and the interpretation of time consistency (based on time perception) given by the subjective discounted utility model. This result may be alternatively presented with a more traditional means for intertemporal choices called Prelec's criterion. The Prelec's criterion $\gamma_0(t)$ is a useful measure of the nature of impatience for all discount functions[8]. For convenience, we make the assumption for the following result that decision weights m are twice differentiable in argument t.

Proposition 5.6

In the SDU model, he DM is time consistent if and only if the Prelec's criterion for any decision period i *is equal to the Prelec's criterion at the initial period*

$$TC \Leftrightarrow \gamma_i(t) = -\rho''(t)/\rho'(t)$$

Proposition 5.6 says that for hyperbolic agents, the decreasing impatience rate for time consistent hyperbolic Decision makers should be independent from the decision period. As a consequence, The Prelec's criterion might be used to dissociate two kinds of inconsistent decision makers: first, DM whose decreasing impatience rates (for given t) are increasing function of decision period – namely $i < j \Longrightarrow \gamma_i(t) < \gamma_j(t)$ – and second, DM

[8] Prelec (2004) proposes that the convexity of the log of the discount function $f(t) = \ln \mu(t)$ is a measure of the impatience degree with $\gamma_0(t) = -f''(t)/f'(t)$.

whose decreasing impatience rates (for given t) are decreasing function of decision period – namely $i < j \Longrightarrow \gamma_i(t) > \gamma_j(t)$. The first category of inconsistent decision makers reveals "classical" time inconsistency such as procrastination or temptation. The second category of decision makers induces an interesting but marginal studied case for which time inconsistency rises, for instance, when the DM exhibits perpetual sacrifice (the opposite case to procrastination) and never enjoys the pleasant task.

Section 6 Concluding Remarks

The intertemporal decision theory assumes a clear division of tasks between the exponential discounting, used to fit normative time consistent behavior, and hyperbolic discount functions, dedicated to observed time inconsistent behaviors. The Subjective Discounted Utility fully explains discounting by the nature of the DM's time perception, and as a consequence, time consistency can be interpreted with time distance properties. In this way, hyperbolic decision makers may be time consistent. Thus, time consistency has not exclusive concern with special discount function, but rather, is a matter of time perception transmission. This paper concludes by a claim in favor of a larger consideration for time perception in intertemporal choices and for the critical issue of time consistency, especially for long term perspectives as in retirement or environmental issues for which the time perception of a DM is likely to change. This claim is reinforced by considering the various fields of application connected to the literature on time consistency/inconsistency issue.

Appendices

Theorem 2.1

A1. $\forall i \in \mathcal{P}, \succsim_i$ is a regular binary relation if and only if $\succsim_i$ satisfies the three following axioms and the existence of a neutral element.

A11. $\forall i \in \mathcal{P}, \succsim_i$ is a weak order (complete and transitive binary relation)

A12. $\forall i \in \mathcal{P}, \succsim_i$ is monotonic

$$\forall k \in S, c_k \geq c'_k \text{ and } \exists k \in S, c_k > c'_k \text{ then } (c;\tau) \succ_i (c';\tau)$$

A13. $\forall i \in \mathcal{P}, \succsim_i$ is continuous

$\forall k \in S, \exists N_k > 0$ such that for all δ_k, $0 < \delta_k < N_k$, we have:

$$(c;\tau) \succ_i (c';\tau') \Longrightarrow (c;\tau) \succ_i (c'_0, \dots, c'_k, \dots; \tau'_0, \dots, \tau'_k + \delta_k, \dots)$$
$$(c;\tau) \succ_i (c';\tau') \Longrightarrow (c;\tau) \succ_i (c'_0, \dots, c'_k + \delta_k, \dots; \tau'_0, \dots, \tau'_k, \dots)$$

A11 and A12 are very standard axioms and insure the existence of a representative intertemporal utility function. Observe that continuity holds for both payments and periods. First Order Separability (A2a) was first suggested by Koopmans (1972) and by Burness (1976). The utility representation (1) is directly obtained from Harvey (1995). $\forall i \in \mathcal{P}, \succsim_i$ satisfies A1-A2a if and only if there exists an intertemporal utility function $U_i : \Delta_i^* \rightarrow \mathbb{R}$ such that:

$$(c;\tau) \succsim_i (c';\tau') \Leftrightarrow U_i(c;\tau) \geq U_i(c';\tau')$$
$$\forall (c;\tau) \in \Delta^*{}_i, U_i(c;\tau) = \sum_{k=0}^{+\infty} v_i(c_k;\tau_k)$$

Function v_i is continuous and strictly increasing with c_k for all fixed $\tau_k \in \mathcal{P}$, it is continuous with τ_k for all fixed $c_k \in \mathcal{C}$ and $v_i(c^*{}_k;\tau_k) = 0$. The intertemporal utility function U_i is unique up to a positive transformation.

However, we cannot directly implement Harvey theorem (page 325) for our representative function since the latter has been built for only one decision period. That is why Second Order Separability was slightly amended to permit decision period independence. To understand the effect of A2b on the representation of time preferences, let us introduce the following lemma on multipliable separability for multi variable functions.

Lemma 1. Separability for a multi variable function

Given f a function defined on the set T^n, $n \in \mathbb{N}$, g defined on T^{n-1}, and h on T, $\forall a_1, a_2, \dots a_n, b_1, b_2, \dots b_n \in T$,

$$\frac{f(a_1, a_2, \dots, a_{n-1}, a_n)}{f(a_1, a_2, \dots, a_{n-1}, b_n)} = \frac{f(b_1, b_2, \dots, b_{n-1}, a_n)}{f(b_1, b_2, \dots, b_{n-1}, b_n)} \Leftrightarrow f(a_1, a_2, \dots a_n) = g(a_1, a_2, \dots a_{n-1}) \times h(a_n)$$

($\Leftarrow$) is immediate

($\Rightarrow$) If $f(a_1, a_2, \dots, a_{n-1}, a_n)/f(a_1, a_2, \dots, a_{n-1}, b_n) = f(b_1, b_2, \dots, b_{n-1}, a_n)/f(b_1, b_2, \dots, b_{n-1}, b_n)$ for all $a_1, a_2, \dots a_n, b_1, b_2, \dots, b_n \in T$ and setting $b_1 = b_2 = \cdots = b_{n-1} = b_n$, we have $f(a_1, a_2, \dots, a_{n-1}, a_n) = f(a_1, a_2, \dots, a_{n-1}, b_n)/f(b_n, b_n, \dots, b_n) \times f(b_n, b_n, \dots, b_n, a_n)$.

Fix then $g(a_1, a_2, \dots, a_{n-1}) = f(a_1, a_2, \dots, a_{n-1}, b_n)/f(b_n, b_n, \dots, b_n)$ and $h(a_n) = f(b_n, b_n, \dots, b_n, a_n)$, the desired expression follows.

We finally turn to the proof of our representation theorem. $\forall i, j \in \mathcal{P}$ $(i < j)$, choose x and y such that $(c^*{}_s x; \tau) \sim_i (c^*{}_t y; \tau)$ and $(c^*{}_{s'} x; \tau') \sim_j (c^*{}_{t'} y; \tau')$. Then by (1), $v_i(x, s) = v_i(y, t)$ and $v_j(x, s') = v_j(y, t')$. Moreover take $z \in \mathcal{C}$ such that $(c^*{}_s z; \tau) \sim_i \left(c^*{}_{s,t} x, y; \tau\right)$ and by A2b $(c^*{}_{s'} z; \tau') \sim_j \left(c^*{}_{s',t'} x, y; \tau'\right)$. Hence $v_i(z, s)/v_i(x, s) = v_j(z, s')/v_j(x, s')$ and by lemma 1 $v_i(x, s) = m(i; s) \times u(x)$ with $m(i; s) = v_i(i, s)/v_i(i, i)$ and $u(x) = v_i(x, i)$. Then $m(i; i) = 1$ and by monotonicity $m(i; t) > 0$. Hence by A3, (*) implies $u(0) = 0$ and notice that Harvey (1995) insures that u is unique up to a positive linear transformation and u increases with x induces $u(x) > 0$ for $x > 0$. Lastly, (**) implies that m is decreasing with t and (***) that m is increasing with i. ∎

Observation 3.1

First observe that decision weights may be rewritten as follows: $\forall t \in \mathcal{P}_i, \forall i \in \mathcal{P}$ and for all $k \geq 0$, such that $t = i + k, m(i; t) = m(i; i + k) = m_i(k)$. The weight m_i is defined on $\mathcal{P}$ since $t \geq i \Rightarrow k \geq 0$ and by definition of m_i, continuity of m implies continuity of m_i. Then it is easy to show that m_i simultaneously verifies the three properties of definition 1. Note that $m(i; i) = 1$ implies $m_i(0) = 1$ $(d1)$, $m(i; t) > 0$ gives $m_i(k) > 0$ $(d2)$ and impatience for m induces decreasing m_i with k $(d3)$. Lastly $(d1)$, $(d2)$ and $(d3)$ constrain the image set to $]0; 1]$ and then $m(i; t) = \mu_i(t - i)$ ∎

Proposition 3.2

By (1), $m(i; t) = \mu_0(t)/\mu_0(i)$ directly implies time consistency. So assume time consistency holds and consider two dated rewards (x, s) and (y, t) $\left(s, t \in \mathcal{P}_j, i, j \in \mathcal{P}\right)$ such that $(x; s) \sim_i (y; t)$. Then by (1) we have $m(i; t)/m(i; s) = u(x)/u(y)$ and by time consistency $(x; s) \sim_j (y; t)$ that is $m(j; t)/m(j; s) = u(x)/u(y)$. Setting $j = s$ and $i = 0$ gives $m(j; t) = \mu_0(t)/\mu_0(j)$ ∎

Corollary 3.3

That multiplicative separability induces time consistency is straightforward. Reciprocally, assume time consistency then by proposition 3.2 $m(i; t)/m(i; s) = m(j; t)/m(j; s)$ $\left(s, t \in \mathcal{P}_j, i, j \in \mathcal{P}\right)$. By Lemma 1 we have $m(i; t) = \varphi(i) \times \vartheta(t)$. Moreover consistent weight can be written as $m(i; t) = \mu_0(t)/\mu_0(i)$ and by association $\vartheta(t) = 1/\varphi(t) = \mu_0(t)$. As a consequence, $m(i; t) = m(i; s) \times m(s; t)$. Turn to second part of Corollary 3.3. For all $(x; \tau) \in \Delta_j$ $(i, j \in \mathcal{P}, i < j)$ by (1) we have $U_i(x; t) = \sum_{k=0}^{+\infty} m(i; t_k) u(x_k)$ and $U_j(x; t) = \sum_{k=0}^{+\infty} m(j; t_k) u(x_k)$. Then $U_i(x; t)/m(i; j) = \sum_{k=0}^{+\infty} [m(i; t_k)/m(i; j)] u(x_k)$ and by proposition 3.2 $U_i(x; t)/m(i; j) = \sum_{k=0}^{+\infty} [m(j; t_k)/m(j; j)] u(x_k) = U_j(x; t)$ ∎

Observation 3.4

First assume time consistency holds. Stop Watch Time implies $m(i; i + t) = m(0; t)$ or $m(i; t) = m(0; t - i)$. Time consistency implies $m(i; t) = m(0; t)/m(0; i)$ and by SWT $\mu_0(t - i) = \mu_0(t)/\mu_0(i)$, functional equation whose broadest solution is $\mu_0(t) = \lambda^t$. Conversely, if we retain EDU then $m(0;t) = \lambda^t$ and SWT implies $m(i; i+t) = \lambda^t$ then $m(i;t) = \lambda^{t-i} = m(0;t)/m(0;i)$. ■

Observation 4.1
First note that $m(i;t) = \mu_i(t-i) = \mu\big(f_i(t-i)\big)$. We have to show that $f_i(t-i) = \rho_i(t-i)$ by transposing properties of m to f_i and by verifying that the three properties of definition 2 hold. It is quite easy and the end of the proof is left to reader. ■

Proposition 4.2
The proof focuses on uniqueness. Assume $d(i;t) = \rho_i(t-i)$ is given then setting $i = 0$ associates to $d(0;t)$ a unique correspondent time perception $\rho(t)$. Therefore we have also a unique time premium φ such that $\varphi(i,k,t) = d(i;t) - \rho(t-i)$. ■

Proposition 4.3
First A6 is equivalent to $d(i;t) = d(i+k; t+k+\tau)$ so for $i = 0$ we have $\rho(t) = d(k; t+k+\tau)$ and for $k = i$ the condition becomes $\rho(t) = d(i; t+i+\tau) = \rho_i(t+\tau)$. Obviously, $\rho_i(t+\tau) = \rho(t)$ is a transmission condition. We now determine the connection between φ and τ. By A6 $\big((x,0)\sim_0(y,t) \Rightarrow (x,i)\sim_i(y, i+t+\tau)\big)$ and by (1) we deduce $\rho(t) = d\big(i; i+t+\tau(i;t)\big)$. By transmission condition, $\rho(t) = d(i; i+t) - \varphi(i; i+t)$ then $\varphi(i; i+t) = d(i; i+t) - d\big(i; i+t+\tau(i;t)\big)$ and for $i = 0$, $\varphi(0;t) = \rho(t) - \rho\big(t+\tau(0;t)\big)$. ■

Corollary 4.4
By proposition 4.2, a time distance is uniquely determined by a combination of an initial condition and a transmission condition. Then we only need to focus on the necessary condition. For algebraic distance the proof is immediate. Turn to quasi algebraic distance and start from the equality $d(i;s) = d(i+k; s+k+\tau)$ given by A6. Setting $i = 0$ and by objective time, $d(k; s+k+\tau) = s$ and by the assumption $\tau = k \times (t-i)/(1+i)$, we have $d(k; s+k+sk) = d(k; k+(1+k)s) = s$. With the notation $d(k; k+(1+k)s) = g_k\big((1+k)s\big)$ and setting $t = (1+k)s$, the following equality is obtained $g_k(t) = t/(1+k)$. Last for $k = i$, $d(i; i+t) = t/(1+i)$ or $d(i;t) = (t-i)/(1+i)$. The third part of corollary 4.3 concerning logarithmic time distance function is immediate. ■

Observation 4.5
Assume time preservation holds then for $t \in \mathcal{P}_i$ $(i \in \mathcal{P})$, $d(i;t) = d(0; t-i)$. Hence by (1) and observation 4.1 $m(i;t) = m(0; t-i)$ must hold and SWT is verified. The demonstration for the reciprocal implication will be proved by the same way. ■

Theorem 4.6
By Observation 4.1 $m(i;t) = \mu \circ d(i;t)$ and by Proposition 3.2 time consistency holds if and only if $m(i;t) = \mu_0(t)/\mu_0(i)$. As a consequence the DM is time consistent iff $\mu \circ d(i;t) = \mu_0(t)/\mu_0(i)$ where μ and μ_0 are two discount functions not necessarily equal. Observe that as μ is continuous and strictly decreasing over $\mathcal{P}$, then μ^{-1} does exist. The remaining demonstration will prove that the function $\pi(i;t) = \mu^{-1}(\mu_0(t)/\mu_0(i))$ is a time distance function. By $(d1)$ in definition 1, $\mu^{-1} : \,]0;1] \to [0;+\infty[$. As μ satisfies $(d2)$ then $\mu_0(t)/\mu_0(i) > 0$ and by impatience $t \geq i$ implies $\mu_0(t)/\mu_0(i) \leq 1$. As a result $\pi : \mathcal{P} \times \mathcal{P}_i \to [0;+\infty[$. Furthermore μ is continuous over $\mathcal{P} = [0;+\infty[$ so over $\mathcal{P}_i \subset \mathcal{P}$, μ^{-1} is continuous over $]0;1]$ therefore π is continuous over $\mathcal{P} \times \mathcal{P}_i$. By definition of μ^{-1}, we have $\pi(i;t) \geq 0$ $(p1)$ and π is increasing with t since $\uparrow t \Rightarrow \mu_0(t) \downarrow \Rightarrow \mu_0(t)/\mu_0(i) \downarrow \Rightarrow \mu^{-1}(\mu_0(t)/\mu_0(i)) \uparrow$ and decreasing with i for $\uparrow i \Rightarrow \mu_0(i) \downarrow \Rightarrow \mu_0(t)/\mu_0(i) \uparrow \Rightarrow \mu^{-1}(\mu_0(t)/\mu_0(i)) \downarrow$ $(p2)$. Lastly $\pi(i;i) = \mu^{-1}(\mu_0(i)/\mu_0(i)) = \mu^{-1}(1) = 0$ $(p3)$. Observe that time distance $d(i;t) = \mu^{-1}(\mu_0(t)/\mu_0(i))$ is not unique since no constraint is imposed on μ. ■

Corollary 4.7
The proof is straightforward.

Theorem 5.1
To obtain $\mu\big(\rho(t)\big) = \lambda^t$ we restate a theorem by Lapied and Renault (2012). According to Theorem 2.1, axioms A.1 to A.3 induce the functional representation (3). Therefore if $(x,s) \sim_0 (y,t)$ then A7 implies for all $s,t,k \in \mathcal{P}$, $\mu\big(\rho(t)\big) \times \mu\big(\rho(s)+\rho(k)\big) = \mu\big(\rho(s)\big) \times \mu\big(\rho(t)+\rho(k)\big)$. Hence choosing $\rho(k) = \rho(t) - \rho(s)$, we obtain $\ln\Big(\mu\big(\rho(t)\big)\Big) = \frac{1}{2}\ln\Big(\mu\big(\rho(t)-\rho(k)\big)\Big) + \frac{1}{2}\ln\Big(\mu\big(\rho(t)+\rho(k)\big)\Big)$. Since $f(t) = \ln\Big(\mu\big(\rho(t)\big)\Big)$ is continuous then

the previous functional equation is a Jensen's equation whose broadest solution is $\ln\Big(\mu\big(\rho(t)\big)\Big) = a\rho(t) + b$. As $\mu(0) = 1$ and $\rho(0) = 0$ so $b = 0$. Hence $\ln\big(\mu(t)\big) = a\rho(t)$ with $a < 0$ since μ is a discount function. Finally we obtain $\mu(j) = \lambda^{\rho(j)}$ $(\lambda = e^{-r}, r = -a)$.
Conversely, the SDU verifies (3) so according to Theorem 2.1, A.1 to A.3 are implied. In addition, if for all $t, k \in \mathcal{P}$ $\mu(t) = \lambda^{\rho(t)}$ hence $\mu\big(\rho(t) + \rho(k)\big) = \lambda^{\rho(t)+\rho(k)}$ then under (3) we have $(x, s) \sim_0 (y, t) \Rightarrow \Big(x, \rho^{-1}\big(\rho(s) + \rho(k)\big)\Big) \sim_0 \Big(y, \rho^{-1}\big(\rho(t) + \rho(k)\big)\Big)$. Since $\mu(t) = \lambda^t$, then by identification, $\mu\big(d(i;t)\big) = \lambda^{d(i;t)}$ immediately follows. ■

Corollary 5.2
The proof of Corollary 5.2 is immediate. However, to clarify the core meaning of this corollary, assume that decision weights m are twice differentiable in t. Then Corollary 5.2 is illustrated by the use of the Prelec's criterion (2004). Recall that the Prelec's criterion is measure of decreasing impatience for all discount functions. We show that the Prelec's criterion in the SDU model only depends on time distance derivatives. The Dynamic Prelec's criterion for decision period i is $\gamma_i(t) = -\frac{\partial^2 m}{\partial t^2}(i;t) / \frac{\partial m}{\partial t}(i;t) = -m_2^{''}(i;t)/m_2^{'}(i;t)$. By theorem 5.1, the Prelec's criterion in the SDU model is reduced to the simpler following expression $\gamma_i(t) = -d_2^{''}(i;t)/d_2^{'}(i;t)$. Then the nature of impatience is integrally determined by the time distance properties. ■

Observation 5.3
For $s, t \in \mathcal{P}_i, s < t$ and $i \in \mathcal{P}$, the triangular equality $d(i;t) = d(i;s) + d(s;t)$ is a Sincov functional equation and by Aczél (1966), the broadest solution of such an equation is $d(i;t) = h(t) - h(i) = h(t) - h(0) - h(t) + h(0) = \rho(t) - \rho(i)$. For the triangular inequality, $d(i;t) > d(i;s) + d(s;t)$ implies $d(s;t) < d(i;t) - d(i;s)$ and for $i = 0$ we have $d(s;t) < \rho(t) - \rho(s)$. The reversed triangular inequality implies the opposite inequality. ■

Proposition 5.4
By proposition 3.2, time consistency implies $m(i;t) = \mu_0(t)/\mu_0(i)$ which can be rewritten in the SDU approach as $\lambda^{d(i;t)} = \lambda^{\rho(t)-\rho(i)}$ and the triangular equality holds. The reciprocal implication is straightforward. ■

Proposition 5.5
Set $(x, s) \sim_0 (y, t)$ then by the initial axiom $(x, s + k) \sim_0 (y, t + k + \theta)$ and by the transmission axiom $(x, s + k) \sim_k (y, t + k + \tau)$.
Assume time consistency then $(x, s + k) \sim_0 (y, t + k + \theta) \Rightarrow (x, s + k) \sim_k (y, t + k + \theta)$ hence $(y, t + k + \tau) \sim_k (y, t + k + \theta)$ and by impatience $\tau = \theta$. Reciprocally, assume $\tau = \theta$. Setting $s + k = l$ and $t + k + \theta = l'$ gives $(x, l) \sim_0 (y, l') \Rightarrow (x, l) \sim_k (y, l')$. ■

Proposition 5.6
The proof follows from Corollary 3.3, Proposition 5.2 and Observation 5.3. Assume first $\gamma_i(t) = \gamma_0(t)$ then by proposition 4 Time consistency holds. Now assume time consistency then by proposition 5.2 $d(i;t) = \rho(t) - \rho(i)$ and by observation 5.1, $\gamma_i(t) = -d_2^{''}(i;t)/d_2^{'}(i;t)$ and $\gamma_0(t) = -\rho^{''}(t)/\rho^{'}(i;t)$. Therefore $\gamma_i(t) = -\rho^{''}(t)/\rho^{'}(t) = \gamma_0(t)$. ■

References

Akerlof G., (1991). Procrastination and Obedience. *American economic Review 81*, 1-19.

Albrecht M. and M. Weber (1995). Hyperbolic Discounting Models in Prescriptive Theory of Intertemporal Choice. *Zeitschrift Fur Wirtschafts-U Sozialwissenschaften 115*, 535-568.

Aczél, J. (1966). Lectures on Functional Equations and Their Applications. *Academic Press, NY.*

Becker, G. and K. M. Murphy, (1988). A theory of rational addiction. *Journal of Political Economics 96* (4), 675-701.

Bleichrodt, H, K.I.M. Rohde and P.P. Wakker, (2009). Non-Hyperbolic Time Inconsistency. *Games and Economic Behavior 66*, 27-38.

Burness, H. S., (1976). On the Role of Separability Assumptions in Determining Impatience Implications. *Econometrica 44*, 67-78.

Chateauneuf A. and Y. Rébillé (2004). Some characterizations of non-additive multi-period models. *Mathematical social sciences 48*, 235-250.

De Waegenaere A. and P. Wakker (1999). Nonmonotonic Choquet Integrals. *Journal of Mathematical Economics 36*, 45-60.

Drouhin N., (2009). Hyperbolic discounting may be time consistent. *Economics Bulletin 29* (4), 2549-2555.

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

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

Gilboa, I. (1989). Expectation and variation in Multi-period decisions. *Econometrica 57* (5), 1153-1169.

Harris C. and D. Laibson (2001). Dynamic Choices of Hyperbolic Consumers. *Econometrica 69* (4), 935-957.

Harvey C., (1994). The reasonableness of non-constant discounting. *Journal of Public Economics 53*, 381-399.

Harvey C., (1995). Proportional Discounting of future Costs and Benefits. *Mathematics of Operations Research 20* (2), 381-399.

Koopmans T.C., (1960). Stationary ordinal Utility and Impatience. *Econometrica 28* (2), 287-309.

Koopmans, T. C. (1972). Representation of preference orderings over time. In McGUIRE and RADNER (Eds.), *Decision and Organization, a Volume in honor of Jacob Marschak.*, pp. 79-100.

Laibson D., (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*, 443–477.

Lapied, A. and O. Renault (2012). A Subjective Discounted Utility Model. *Working paper*.

Loewenstein G. F. and R. H. Thaler (1989). Anomalies: Intertemporal Choice. *The Journal of Economic Perspectives 3* (4), 181-193.

Loewenstein, G. F. and D. Prelec (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics 107*, 573–597.

Loewenstein, G. F. and D. Prelec (1997). Beyond Time Discounting. *Marketing Letters 8* (1), 97-108.

Phelps E.R and R.A. Pollak (1968). On Second-Best National Saving and Game Equilibrium Growth. *The Review of Economic Studies 35* (2), 185-199.

Prelec D., (2004). Decreasing impatience: A criterion for non-stationary time preference and 'hyperbolic' discounting. *Scandinavian Journal of Economics 106*, 511–532.

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

Strotz R., (1956). Myopia and inconsistency in dynamic utility maximisation. *Review of Economic Studies 23* (2), 165-180.

Thaler, R. H., (1981). Some Empirical Evidence on Dynamic Inconsistency. *Economic Letters 8*, 201-207.

Zauberman, G., K. Kim, S. Malkoc and J. Bettman (2009). Discounting Time and Time Discounting: Subjective Time Perception and Intertemporal Preferences. *Journal of Marketing Research 46* (4), 543-556.