



IZA DP No. 1228

## A Vague Theory of Choice over Time

Paola Manzini  
Marco Mariotti

July 2004

# A Vague Theory of Choice over Time

**Paola Manzini**

*Queen Mary, University of London  
and IZA Bonn*

**Marco Mariotti**

*Queen Mary, University of London*

Discussion Paper No. 1228  
July 2004

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
Email: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### A Vague Theory of Choice over Time\*

We propose a novel approach to modelling time preferences, based on a cognitive shortcoming of human decision makers: the perception of future events becomes increasingly ‘blurred’ as the events are pushed further in time. We axiomatise a class of preference representations which can be specialised to rationalise ‘anomalies’ such as preference reversals and cyclical preferences. We also present an application to a form of time inconsistency which we dub ‘the now or never fallacy’.

JEL Classification: A12, C70, D90

Keywords: time preferences, hyperbolic discounting, preference reversal, cycles, intransitive preferences

Corresponding author:

Paolo Manzini  
Department of Economics  
Queen Mary  
University of London  
Mile End Road  
London E 1 4 NS  
United Kingdom  
Email: [p.manzini@qmul.ac.uk](mailto:p.manzini@qmul.ac.uk)

---

\* Part of the work for this paper was carried out while we were visiting the Institute for Economic Analysis at Universitat Autònoma de Barcelona. We are grateful for their generous hospitality. We wish to thank E. Diamantoudi, P. Jehiel, G. Herringer, M. Mandler, C. Ponsati, V. Skreta, J. Valimaki and X. Vilà for comments, and audiences at Bocconi, Catholic University Milan, CORE, University College London, University of Newcastle, University of Pisa, Rice University, Royal Holloway University of London, Tel Aviv University, Universitat Autònoma de Barcelona, SGTM in Seville and IMGTA in Urbino. The usual disclaimers apply.

# 1 Introduction

Most economic decisions involve an appraisal of future outcomes. The standard model used to deal with this issue is the exponential discounting model (EDM), according to which<sup>1</sup> an outcome  $x$  available at time  $t$  is evaluated as  $\delta^t u(x)$ , with  $\delta$  a constant discount factor and  $u$  the instantaneous utility of  $x$ . However, in the last decade evidence has accumulated contradicting the predictions of the EDM. In particular, it is clear that individuals often exhibit the so-called preference reversal phenomenon, whereby an initial preference between an earlier outcome and a later one is reversed once both outcomes are delayed by the same amount of time.

In the last decade an alternative model, which is consistent with this type of evidence, has enjoyed considerable prominence and success in the literature. This is the model of *hyperbolic discounting* (HDM)<sup>2</sup>, in which the discount factor is a hyperbolic function of time. The central point of this approach is that the rate of time preference between alternatives is not constant but varies, and in particular decreases, with the passage of time. The simplest and most popular form of hyperbolic discounting occurs in a two-parameter model which may be termed the  $\beta - \delta$  model. In this model, the rate of time preference between a present alternative and a future one is  $\beta\delta$ , whereas the rate of time preference between two future alternatives is  $\delta$ . So we may have for example  $u(x) > \beta\delta u(y)$  and  $\delta^{t+1}u(y) > \delta^t u(x)$ , ‘rationalising’ a preference reversal. The HDM approach is extremely fruitful in addressing some of the issues raised by the experimental evidence, so much so that it has almost become a new consensus.

Yet we believe that some closer scrutiny, and an appraisal of the cognitive sources of the ‘anomalies’ is needed. For example Rubinstein ([25] and [26]) shows experimentally that precisely the same type of decision situations that may create a difficulty for the EDM may also be problematic for the HDM. Similarly, Frederick *et al.* [8] doubt that the change from a constant to hyperbolic functional form to discount future utility is ‘radical’ enough. Furthermore, Read [21] shows that the existing experimental evidence is consistent with other explanations, most notably subadditivity of discounting.

There is also evidence that human decision-makers depart in an even more fundamental way from standard models: they may make *intransitive* choices. Although most data in this direction come from choices under risk, what evidence is available for time preferences suggests that violations of transitivity are even more frequent in this domain. Tversky *et al.* [34] show that a substantial 15% of subjects exhibited cyclical patterns of choice that

---

<sup>1</sup>See Fishburn and Rubinstein [7].

<sup>2</sup>See for instance Phelps and Pollack [18], Loewenstein and Prelec [13], Laibson [12], Frederick *et al.* [8], O’Donoghue and Rabin [15], [16].

could not be explained by ‘framing effects’. Roelofsma and Read’s [22] experiment is even more striking: they found that the *majority* of intertemporal choices were intransitive. This type of evidence, which we consider crucial in guiding our intuition about time preferences, has been somewhat downplayed in the mainstream economic literature on time preferences. This contrasts with the theory of choice under risk, where at least two main approaches (prospect theory and regret theory) explain cyclical choices. A notable exception is the very general paper by Ok and Masatlioglu [17], who have axiomatised a class of models of choice over time. These include as particular cases HDM, Rubinstein’s ([25], [26]) similarity theory and Read’s subadditive discounting among several others, and some of the representations admitted by their axioms are compatible with intransitivities.

We view the body of evidence against the EDM as indicating that human agents use heuristic *procedures* to make choices over time that differ fundamentally from the simple maximisation of a complete preference ordering as assumed both in EDM and in HDM. We suggest that at the root of ‘anomalies’ like preference reversals and intransitivities is a basic cognitive shortcoming, noted by several thinkers on the topic and in popular wisdom: the perception of events that will occur in the future is ‘blurred’, and possibly it becomes increasingly blurred as the events are pushed farther in time<sup>34</sup>. As a consequence, deciding between alternatives to be obtained in the distant future is intrinsically harder than deciding between the same alternatives available earlier. For instance, an individual may find 1000 euros now distinctly preferable to 1100 euros in one year time, but may well hesitate when asked to decide between 1000 euros in ten years or 1100 euros in eleven years. In our view, the evaluation of a time-dependent alternative has, on the one hand, a *pure time preference* aspect: most people prefer alternatives to be available sooner rather than later (this may be a ‘hard-wired’ feature of animal brains, see e.g. Ainslie [3]). On the other hand, the farther two alternatives are in time, the more difficult it is to distinguish between them on the basis of an explicit time-outcome trade-off: we refer to this cognitive condition as *vagueness*. In our theory we propose that behaviour can be rationalised by a combination of three factors:

---

<sup>3</sup>For example, according to Pigou [19] “our telescopic faculty is defective, and we, therefore, see future pleasures, as it were, on a diminished scale”. Also, Böhm-Bawerk [4] stated that “..we limn a more or less incomplete picture of our future wants and especially of the remotely distant ones.”

<sup>4</sup>These notions are well established in proverbs, from “we’ll cross that bridge when we come to it” to “hindsight is better than foresight”, “it will all be the same in a hundred years” for the anglosaxon tradition, to “En cien años todos calvos” and “Tra tre anni beato chi ha un occhio” from the Spanish and Italian tradition, respectively (which translate as “In a hundred years we’ll all be bald” and “Blessed he who still has got an eye left in three years”). Interestingly, for the latter proverb the user is free to change the lenght of time as she sees fit). See for instance Taylor [31], Simpson and Speake [28], Rovira [23], Pittano [20].

- impatience;
- ‘vagueness’ in the perception of future alternatives; and
- a suitable heuristics to make a choice when vagueness is ‘high’.

The decision maker will initially attempt to express a preference between two alternatives based on an explicit trade-off between outcome and time (that is, between his impatience and his desire to obtain more). When unable to do so, he will rely on a ‘vagueness-breaking criterion’. This criterion is based on a lexicographic consideration of the two dimensions (time and outcome) involved in the comparison. That is to say, the decision maker chooses the alternative that is better according to the dimension that he regards as the most important. This is in accord with evidence and theory in the psychology literature. As for example Tversky *et al.* [35] note,

‘Because it is often unclear how to trade one attribute against another, a common procedure for resolving conflict in such situations is to select the option that is superior on the more important attribute. This procedure, which is essentially lexicographic, has two attractive features. First, *it does not require the decision maker to assess the trade-off between the attributes*, thereby reducing mental effort and cognitive strain. Second, it provides a compelling argument for choice that can be used to justify the decision to oneself as well as to others.’ (p. 505, our italics)

The economic literature on individual preferences traditionally looks for representation results that ‘compress’ preferences within a single ‘indicator’, a utility function which summarises the agent’s evaluation of all aspects of an alternative. On the contrary, the psychological theories of choice tend to be completely heuristic based. We see our theory as ‘spanning’ the gap between the two approaches since, although agents’ evaluate alternatives ‘in their entirety’, they are not always able to tell two alternatives apart: it is only in the presence of vagueness that they rely on heuristic procedures.

Among heuristics based theories the closest to our approach is Tversky’s [33] ‘*lexicographic semiorder*’, according to which agents rely on their ranking of the attributes of a (time dependent) alternative in a lexicographic way when choosing between different alternatives. The first attribute of each alternative is compared, and if the difference exceeds some threshold value then a choice is made accordingly. If the threshold value is not exceeded, then the agent compares the second attribute of each alternative, and so on.

Both this approach and our theory are close in spirit to Rubinstein's [26] suggestion to extend his own similarity-based approach from risk to time preferences. He argues that, when evaluating alternatives, agents may perceive these alternatives to be 'similar' in either the time or the outcome dimension. At the formal level, this notion of similarity is closely related to our notion of vagueness. The crucial difference is that in our model vagueness applies to alternatives (i.e. outcome-dates pairs), not only to each date or outcome dimension separately. There are two aspects to this. First, vagueness (or similarity) in one dimension may not be defined independently of the other dimension. For example, 1 Million today may be clearly distinct from 1.01 Million today; but from today's perspective the two amounts in thirty years time may well be considered similar<sup>5</sup>. Second, 'two-dimensional' similarities between alternatives (outcome-date pairs) may be independent of similarities along each dimension: you may be able to distinguish between 1 Million and 1.01 Million, and also between the dates '1 year from now' now and '2 years from now', yet be unable to express a preference between 1.01 Million 2 years from now and 1 Million 1 year from now. These differences are important. Indeed, Rubinstein [24] shows that his similarity approach generates an essentially unique transitive preference relation on alternatives<sup>6</sup>, and takes it as evidence that utility theory is hardly reconcilable with experimental data. On the contrary, as observed above, our theory, though utility based, can easily explain a range of both transitive and cyclical choices.

Our contribution is two-fold. First, we provide a class of parsimonious representations of an individual's basic time-outcome trade-off, in which a standard utility function (of time and outcomes) is combined, in an additive way, with a 'vagueness function'. At the formal level, these representations take the form of interval orders or semiorders. Second, we focus on a very simple such representation - which we dub the  $\sigma - \delta$  model - and combine it with the 'vagueness-breaking' criterion described above in order to examine choice behaviour. Needless to say, heuristics by definition cannot be axiomatised; we show how a simple 'rule of thumb' which has been widely studied in the psychological literature can be used to complete the basic transitive partial order when the latter leaves the decision maker 'vague'. In this way we can explain for instance the phenomenon of preference reversal with an interpretation that seems to us more adherent to the psychological basis for it. The model can also account for *cyclical* patterns of choice, which are obviously incompatible with theories such as HDM and EDM, based as they are upon the maximisation of a preference ordering. Moreover, our framework can address a type

---

<sup>5</sup>This aspect has been underlined and studied, in a risk context, by Aizpurúa *et al.* ([2] and [1]) and Uriarte [36], who develop the concept of 'correlated similarities'.

<sup>6</sup>Vilà [38] shows that with three dimensions Rubinstein's similarity approach generates intransitivities. See also Gilboa and Lapson [9].

of time inconsistency that so far has not received attention, and which we dub the ‘now or never fallacy’: an agent engages in some activity prematurely for fear that he will keep delaying it to an undesirable extent. The now or never fallacy hinges directly on a cyclical pattern of preference (as well as time-inconsistency). Although it accords with introspection, neither HDM nor EDM could explain it.

The structure of the paper is as follows. We start by postulating a partial order (the ‘primary criterion’) on the set of outcome-date pairs, together with a ‘secondary criterion’ which embodies the vagueness-breaking heuristics (section 2). Next, we narrow down further this partial order, and impose axioms which specialise it to an interval order or a semiorder, which can be represented in a standard, additively separable, way (section 3). In section 4 we present our  $\sigma - \delta$  model and consider preference reversal, cycles and the now or never fallacy application. Section 5 concludes.

## 2 The Model

In this section we present a general model which includes, as a special case, the simple  $\sigma - \delta$  formulation studied later. Let  $X$  indicate a set of outcomes (e.g. money amounts) on which there exists a reflexive, complete and transitive preference relation  $R$ . We denote the symmetric and asymmetric components of  $R$  by  $I$  and  $P$ , respectively. Let  $T$  be the set of dates. The set of alternatives is a set of outcome-date pairs  $A \subseteq X \times T$ .

We imagine that the preference relation  $\succeq^*$  of the individual on  $A$  is constructed as follows. There exists a *primary* criterion on the basis of which the individual makes comparisons between alternatives. We interpret the primary criterion as resolving comparisons for which the trade-off between outcome and time involved in the comparison yields, in the perception of the individual, a decisive advantage to one of the alternatives: that is, for example, the loss in the time dimension of one alternative clearly outweighs its advantage in the outcome dimension. We would also expect a primary criterion to operate when there exists a clear relation of ‘dominance’ between alternatives, in the sense that one offers a distinctly better outcome at an earlier date than the other. Because of the cognitive limitations discussed in the introduction, however, there may be pairs of alternatives that cannot be ranked by the primary criterion alone: the outcome-time trade-off does not offer a basis for decision, and the individual has cognitive ‘holes’. In this case, and only in this case, a *secondary* criterion is used. The relation  $\succeq^*$  is derived by a combination of primary and secondary criteria.

More formally, let the primary criterion be a strict partial order  $\succ$  on  $A$  (that is, an irreflexive and transitive binary relation on  $A$ ). When a pair of alternatives  $a, b \in A$  is

not ordered by  $\succ$  we denote this fact by  $a \sim b$ , i.e.

$$(not\ a \succ b, not\ b \succ a) \Leftrightarrow a \sim b$$

and say that  $a$  and  $b$  are *vague*. Under our assumptions the relation  $\sim$  may not be transitive. As mentioned above,  $\sim$  does *not* mean indifference between two alternatives, but simply the inability to express a strict preference. We call this situation *vagueness*.

The secondary criterion is used to rank alternatives which are vague. As explained in the introduction, there is psychological evidence that when individuals are unable to compare alternatives on the basis of a trade-off between the various dimensions of those alternatives, they rely on lexicographic-type comparisons, by focussing on the ‘dominant’ dimension<sup>7</sup>. In our model there are two dimensions, time and outcome. Thus, a natural secondary criterion is to allow the individual to rank two alternatives which are vague under  $\succ$  first according to the preference ordering  $R$  over the (pure) outcomes, and if this still does not result in a strict preference, according to time precedence. Formally, let  $\succeq^*$  denote a complete preference relation (not necessarily transitive) on  $A$ . So we consider the following model:

*Outcome Prominence Model (OPM):* Let  $i = (x_i, t_i) \in A$ . Then:

1.  $a \succ^* b \Leftrightarrow$ 
  - (a)  $a \succ b$  (Primary Criterion), or
  - (b)  $(a \sim b, x_a P x_b)$  or  $(a \sim b, x_a I x_b, t_a < t_b)$  (Secondary Criterion)
2.  $a \sim^* b \Leftrightarrow (a \sim b, x_a I x_b, t_a = t_b)$

The other natural model in place of OPM is:

*Time Prominence Model (TPM):* Let  $i = (x_i, t_i) \in A$ . Then:

1.  $a \succ^* b \Leftrightarrow$ 
  - (a)  $a \succ b$  or
  - (b)  $(a \sim b, t_a < t_b)$  or  $(a \sim b, t_a = t_b, x_a P x_b)$
2.  $a \sim^* b \Leftrightarrow (a \sim b, x_a I x_b, t_a = t_b)$

---

<sup>7</sup>See e.g. Slovic [29], Tversky *et al.* [35] and Shafir *et al.* [27].

TPM would yield results which are qualitatively similar to the ones we derive in the paper with OPM. However TPM - in our  $\sigma - \delta$  specialization - would lead to a clash with a large body of experimental evidence which on the contrary is compatible with OPM. In particular, while we can explain ‘preference reversals’ with both models, those obtained with TPM are in a direction opposite to the one which is normally observed in practice. For this reason we will focus on OPM.

### 3 General representation results

In this section we provide a number of representation results for the primary ordering  $\succ$ . These are based on various assumptions on both the structure of the set of alternatives and the cognitive abilities of the decision maker. All the results represent preferences by means of two real valued functions  $u$  and  $\sigma$  on the set of alternatives.

The natural interpretation is that the  $u$  function expresses the ‘pure’ time preferences of the individual. This function can be chosen to be monotonic in the natural direction in both outcome and time. On the other hand  $\sigma$  is the ‘vagueness function’ which captures the imaginative difficulty the individual faces when appraising outcomes in the future. There are two notable aspects of the representations. First, the vagueness term enters additively. Second, the vagueness when appraising whether  $a = (x, t_x)$  is better than  $b = (y, t_y)$  is a function of the components of  $b$  alone; it does not depend on  $t_x$  or  $x$ .

In what follows we say that a preference ordering  $\succ$  is  $(u, \sigma)$ -representable if there exist real valued functions  $u$  and  $\sigma$  on  $A$  such that

$$(x, t_x) \succ (y, t_y) \Leftrightarrow u(x, t_x) > u(y, t_y) + \sigma(y, t_y) \quad \forall (x, t_x), (y, t_y) \in A$$

It is worth comparing our class of representations with that of Ok and Masatlioglu [17]. They consider a *complete* binary preference relation  $B$  over a set of outcome-date pairs and axiomatise the following representation class:  $(x, t) B (y, s)$  if and only if  $U(x) \geq U(y) + \varphi(s, t)$ , where  $U$  is interpreted as an instantaneous utility function and  $\varphi$  captures the effect of time delay. In this representation, which is not an interval order, the ‘contributions’ of outcome and time to the agent’s utility are separated, unlike in our model. Note that in Ok and Masatlioglu’s approach cycles can be accounted for without resorting to a secondary criterion. Our view is different: the primary criterion represents the ‘fully rational’ component of decision making, hence we assume it transitive. In our approach intransitivities are the byproduct of resorting to the rule of thumb invoked when ‘full rationality’ is not decisive.

### 3.1 Finite sets of alternatives

With vague time preferences, an individual is not always able to compare two alternatives  $a$  and  $b$ . However we require that the individual has at least some discriminatory ability regarding the sets of alternative *dominated* by  $a$  and  $b$ , respectively. Formally, let  $L(a)$  denote the lower contour set of  $a \in A$ , that is  $L(a) = \{b | a \succ b\}$ , and let  $V(a)$  denote the vague set of  $a$ , that is  $V(a) = \{b | a \sim b\}$ . Then we assume:

**Discrimination:** For every alternatives  $a, b \in A$  either  $L(a) \subseteq L(b)$  or  $L(b) \subseteq L(a)$ .

Transitivity of  $\succ$  ensures that  $L(b) \subset L(a)$  whenever  $a \succ b$ , thus Discrimination has a bite only in the case of vague alternatives. Two alternatives  $a$  and  $b$  may be so vague that they also have the same set of worse alternatives. But suppose that the agent perceives that alternatives that are not worse than  $a$  are worse than  $b$ . In this case, the agent perceives some kind of *improvement* in moving from  $a$  to  $b$  (though not strong enough to make  $b$  preferred to  $a$ ). The property of Discrimination requires that after the ‘improving’ move to  $b$  no alternative that was dominated by  $a$  ceases to be dominated (and it poses the converse requirement if the move to  $b$  was perceived to be negative). In other words, moving from one alternative to the other may involve either an (weak) improvement or a (weak) worsening in the above sense, but not both.<sup>8</sup>

Next, we consider some straightforward monotonicity properties:

**Time Monotonicity (i):** For every  $x \in X$  and  $t_1, t_2 \in T$  with  $t_2 > t_1$ :  $L(x, t_2) \subseteq L(x, t_1)$ .

**Time Monotonicity (ii):** For every  $x, y \in X$  and  $t, t_1, t_2 \in T$  with  $t_2 > t_1$ :  $(x, t_2) \in V(y, t)$  implies  $(x, t_1) \notin L(y, t)$ .

The requirements of Time Monotonicity are that: first, anything which is dominated by some alternative  $a$  must also be dominated by an alternative with the same outcome as  $a$  but at an earlier time; and, second, that given a relationship of vagueness between two contemporaneous alternatives, it cannot be that anticipating one of them makes it dominated by the other one. Note that these allow for two alternatives with the same outcome available at different times to be vague. Similar requirements can be imposed in the outcome dimension:

---

<sup>8</sup>An alternative justification for Discrimination is the ‘sure-thing’ property if alternatives were allowed to have a risk dimension as well, i.e. lotteries of the type  $(\frac{1}{2}, a; \frac{1}{2}, b)$  over the riskless alternatives  $a$  and  $b$  were included in the domain. Sure-thing implies that if  $a \succ b$  and  $c \succ d$  then  $(\frac{1}{2}, a; \frac{1}{2}, c) \succ (\frac{1}{2}, b; \frac{1}{2}, d)$  and similarly for  $\sim$ . It is easily seen that Discrimination on the riskless alternative would follow.

**Outcome Monotonicity (i):** For every  $x_1, x_2 \in X$  and  $t \in T$  with  $x_2 P x_1$ :  $L(x_1, t) \subseteq L(x_2, t)$ .

**Outcome Monotonicity (ii):** For every  $y, x_1, x_2 \in X$  and  $t, t' \in T$  with  $x_2 P x_1$ :  $(x_1, t) \in V(y, t')$  implies  $(x_2, t) \notin L(y, t')$ .

Our next representation result applies to the case where the number of alternatives is arbitrarily large but finite.

**Proposition 1** *Let Discrimination, Time Monotonicity (i) and Outcome Monotonicity (i) hold. Let  $L(a)$  be finite for all  $a \in A$ . Then  $\succ$  is  $(u, \sigma)$ -representable with  $\sigma(a) > 0$  for all  $a \in A$  and  $u$  non-decreasing in outcome and non-increasing in time.*

**Proof.** For any  $a \in A$ , define  $u(a) = \#L(a)$ . By Time Monotonicity (i) and Outcome Monotonicity (i), the function  $u$  is weakly monotonic in the desired direction with respect to time and outcomes. Next, define  $\sigma(a)$  as follows. Choose

$$a^* \in \arg \max_{b \in V(a)} \#L(b)$$

Then let

$$\sigma(a) = \max \left\{ \frac{1}{2}, u(a^*) - u(a) \right\}$$

Observe that for any two alternatives  $c, d$  with  $c \succ d$  it must be  $u(c) > u(d)$  (since  $L(d) \subseteq L(c)$  by the transitivity of  $\succ$ , and since  $d \in L(c)$  but  $c \notin L(c)$ , implying  $\#L(c) > \#L(d)$ ).

We show that  $u$  and  $\sigma$  as defined above represent  $\succ$ .

Let  $a \succ b$ . Then  $u(a) > u(b)$  by the above observation. Now suppose in negation that

$$u(a) \leq u(b) + \sigma(b)$$

By definition  $\sigma(b) = \max \left\{ \frac{1}{2}, u(b^*) - u(b) \right\}$ . If  $\sigma(b) = \frac{1}{2}$  we get an immediate contradiction (since  $u(a) - u(b) > \frac{1}{2}$ ). So let  $\sigma(b) = u(b^*) - u(b)$  and therefore

$$u(a) \leq u(b) + \sigma(b) = u(b^*)$$

By the above inequality and the definition of  $u$ ,  $\#L(b^*) \geq \#L(a)$ . Then by Discrimination  $L(b^*) \supseteq L(a)$ . By the fact that  $b \in L(a)$  this implies  $b \in L(b^*)$ , contradicting  $b \in V(b^*)$ .

Consider now  $a \sim b$ . We need to show that both  $u(a) \leq u(b) + \sigma(b)$  and  $u(b) \leq u(a) + \sigma(a)$ . If  $b = a^*$  then  $u(b) = u(a^*) \leq u(a) + \sigma(a)$ , and if  $b \neq a^*$  then, by the definition of  $a^*$ ,  $u(b) \leq u(a^*) \leq u(a) + \sigma(a)$ . Therefore in all cases  $u(b) \leq u(a) + \sigma(a)$ . The same argument applies interchanging  $a$  and  $b$ , so that  $u$  and  $\sigma$  as defined above represent  $\sim$ .

Finally assume  $u(a) > u(b) + \sigma(b)$ . Then by Discrimination  $L(a) \supset L(b)$ . So it cannot be  $b \succ a$ , since then we would have  $a \in L(b)$  but  $a \notin L(a)$ , a contradiction. It cannot be  $b \in V(a)$  either, since then  $u(a) \leq u(b) + \sigma(b)$ . ■

The case of a finite set of alternatives  $A$  is a natural setting for this result. Observe however that all that is required is that  $L(a)$  is finite<sup>9</sup>.

The assumptions of Discrimination and Time and Outcome Monotonicity are too weak to allow for a constant  $\sigma$ , even with simple finite sets of alternatives. We illustrate this with an example (see figure 1, in which the arrows denote dominance and the dashed lines denote vagueness).

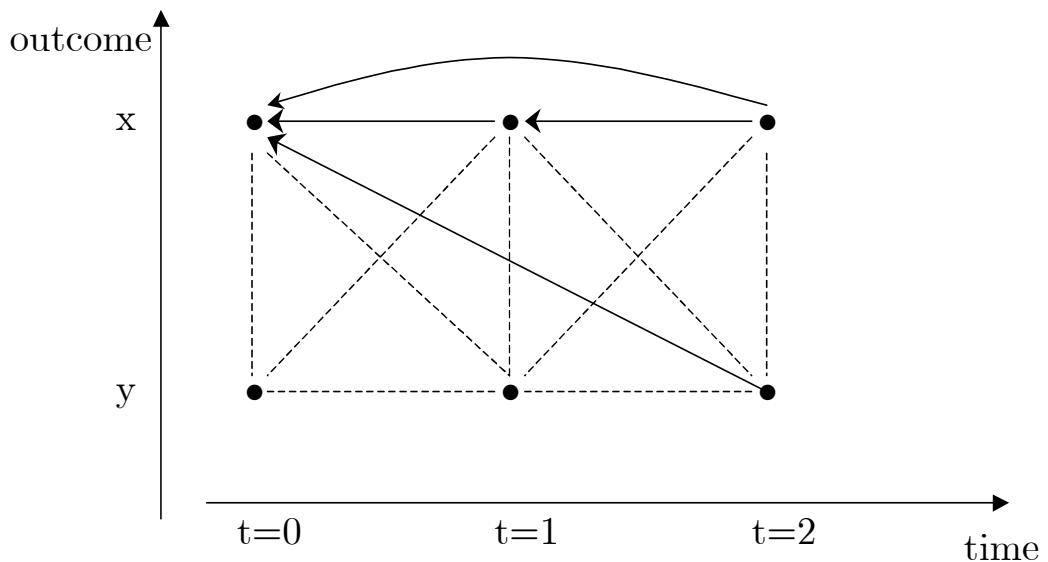


Figure 1: An illustration of example 3.

**Example 2**  $X = \{x, y\}$  with  $xPy$  and  $T = \{0, 1, 2\}$ ,  $A = X \times T$  and  $(x, 0) \succ (x, 1) \succ (x, 2)$ ,  $(x, 0) \succ (y, 2)$  with all other relations not implied by the transitivity of  $\succ$  being ones of vagueness. So,  $L(x, 0) = \{(x, 1), (x, 2), (y, 2)\}$ ,  $L(x, 1) = \{(x, 2)\}$  and for all other

---

<sup>9</sup>This result is related but not the same as Theorem 2.7 in Fishburn [6], who uses a different technique of proof.

$a \in A$ ,  $L(a) = \emptyset$ . Then it is easily checked that these preferences satisfy Discrimination, Time Monotonicity and Outcome Monotonicity, so that there exists a  $u - \sigma$  representation. However, if  $\sigma$  were constant we would have

$$\begin{aligned} u(x, 0) &> u(x, 1) + \sigma \\ u(x, 1) &> u(x, 2) + \sigma \\ u(y, 1) &\leq u(x, 2) + \sigma \\ u(x, 0) &\leq u(y, 1) + \sigma \end{aligned}$$

From the last two inequality it follows that

$$u(x, 0) - \sigma \leq u(x, 2) + \sigma$$

whereas from the first two inequalities we have

$$u(x, 0) - \sigma > u(x, 2) + \sigma$$

a contradiction.

### 3.2 Infinite sets of alternatives

Consider now the case where the set of alternatives is allowed to be infinite, and in particular let  $X = T = [0, 1]$ . A technique similar to the one used for proposition 1 can be applied, provided only that the sets  $L(a)$  are sufficiently regular as to be Lebesgue-measurable (this will be the case in virtually every conceivable application), and that the following hold:

**Continuity:** If  $a = (x, s) \succ b = (y, t)$  then there exists  $\varepsilon > 0$  such that (i)  $(x', s') \succ b$  for all  $x', s' \in [0, 1]$  with  $|x' - x| < \varepsilon$  and  $|s' - s| < \varepsilon$ , and (ii)  $a \succ (y', t')$  for all  $y', t' \in [0, 1]$  with  $|y' - y| < \varepsilon$  and  $|t' - t| < \varepsilon$ .

**Betweenness:** If  $a \succ b$  then there exists  $(x, s) \in L(a) \setminus L(b)$  and  $\varepsilon$  such that  $(x', s') \in L(a) \setminus L(b)$  for all  $x', s' \in [0, 1]$  with  $|x' - x| < \varepsilon$  and  $|s' - s| < \varepsilon$ .

Continuity says that a strict preference between two alternatives should hold for all other alternatives sufficiently close the original two. Betweenness says that there exists a (non-negligible) set of alternatives which lies ‘between’ (in terms of preference) any two strictly ranked alternatives. ‘Between’ in this context means: worse than the better alternative and not worse than the worse alternative. Such assumptions would be satisfied, for example, in a model with ‘thick’ vagueness curves through an alternative, that separate the upper and lower contour sets. However, there may be other patterns of cognitive

'holes', for example one where vagueness occurs very close to an alternative (because the individual does not distinguish such close alternatives), it occurs again at far away alternatives (for example because of the passage of time), and does not occur in the intermediate region.

**Proposition 3** *Let  $X = T = [0, 1]$ . Let Discrimination, Time Monotonicity (i), Outcome Monotonicity (i), Continuity and Betweenness hold. Then  $\succ$  is  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and  $u$  non-decreasing in outcome and non-increasing in time.*

**Proof.** The proof follows the same steps as the proof of proposition 1, taking care of some additional technical details. It is given for completeness in the Appendix.

We can obtain a similar representation in the case of a countable set of alternatives, using an alternative notion of continuity:

**Proposition 4** *Let  $X$  and  $T$  be countable sets. Let Discrimination, Time Monotonicity (i), Outcome Monotonicity (i) hold. Suppose that  $(\lim \{a_n\}) \in V(a)$  for any sequence  $\{a_n\}$  with  $a_n \in V(a)$ . Then  $\succ$  is  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and  $u$  non-decreasing in outcome and non-increasing in time.*

**Proof.** In the Appendix.

Finally, we provide a result which does not make any structural assumption on the set of alternatives but is formulated directly in terms of a representation.

**Proposition 5** *Let  $\succ$  be  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and let Time Monotonicity (i) and (ii), and Outcome Monotonicity (i) and (ii) hold. If  $\sigma$  can be chosen to be non-increasing in outcome, then  $u$  can be chosen to be non-decreasing in outcome at an arbitrarily large number of alternatives, and if  $\sigma$  can be chosen to be non-decreasing in time then  $u$  can be chosen to be non-increasing in time at an arbitrarily large number of alternatives.*

**Proof.** In the Appendix.

Proposition 5 shows that the utility function can be chosen to display the expected monotonicity properties (on an arbitrarily large number of alternatives) with respect to the passage of time and improvement of outcomes, provided that vagueness does not increase as the outcome improves and it does not decrease as time passes. This latter condition in particular accords well with our motivating intuition, that the passage of time has a blurring effect on the comparisons between alternatives. Of course, this result is not completely satisfactory because we would like to have an assumption on the primitives,

namely the preferences, implying the desired conditions on the function  $\sigma$ . In particular, sufficient conditions for the constancy of  $\sigma$  would be of interest.

In the next section we concentrate on one element of the class of representations obtained, which incorporates the feature of (relative) vagueness increasing with time.

## 4 The $\sigma - \delta$ model

From now on we consider  $X$  as a real interval, e.g. money amounts. For this setting we propose our  $\sigma - \delta$  model, a specialisation of OPM. In doing so we follow the hyperbolic discounting tradition whereby out of the entire hyperbolic family attention has been focused on the application friendly  $\beta - \delta$  version. There are several motives for studying our chosen specialisation. First, it constitutes in a sense the minimal possible departure from the standard EDM; yet its implications are still widely different from that of the EDM. Second, it is parsimonious, in that it depends only on two parameters, which can be potentially estimated (see section 4.4). Third, it has a natural interpretation, and in particular it incorporates the psychological feature at the root of our approach that as choices are pushed further into the future the ‘vagueness’ associated with the outcomes increases.

The  $\sigma - \delta$  model is a representation of the individual’s (primary) preferences that depends on a discount parameter  $\delta \in (0, 1)$  and a vagueness parameter  $\sigma > 0$ . The individual values alternatives using exponential discounting with factor  $\delta$  and linear utility. However, in order to regard an alternative better its discounted utility must exceed that of a competing alternative by at least  $\sigma$  (Primary Criterion). That is, loosely speaking,  $\delta$  measures the way the individual trades off outcomes across time, while  $\sigma$  measures the precision with which such trade-offs can be made (see section 4.4 below for a more precise interpretation). When the individual cannot compare alternatives on the basis of present discounted value, his preferences conform to the Secondary Criterion, that is: he prefers the alternative corresponding to the higher amount of money; if the monetary amounts are the same for both alternatives, then he prefers the alternative corresponding to the earlier date. So in the  $\sigma - \delta$  model  $(x, t_x)$  is chosen over  $(y, t_y)$  whenever either

$$x\delta^{t_x} > y\delta^{t_y} + \sigma$$

or

$$x\delta^{t_x} \leq y\delta^{t_y} + \sigma \text{ and } x > y$$

or

$$x\delta^{t_x} \leq y\delta^{t_y} + \sigma, x = y \text{ and } t_x < t_y$$

## 4.1 Cycles

Several contributions in decision theory have pointed out on both normative and descriptive grounds that human decision procedures may generate non transitivity in choice (see e.g. Tversky [33]; Rubinstein [24]; Tversky et al. [34]; Vilà [38]; Roelofsma and Read's [22]; Read [21]; Ok and Masatlioglu [17]). To understand this phenomenon, consider the following non time-related dietary example. You like your coffee sweet to the point of neglecting dietary considerations, so that one sugar spoon is preferred to half a sugar-spoon. However, the sweetening impact of a quarter spoon is negligible to your tongue: in this case dietary considerations prevail and you prefer half a spoon to three quarters and three quarters to one!<sup>10</sup>

With time preferences, we have argued for an analogous ‘switch of criterion’ when vagueness occurs. So the preference relation  $\succeq^*$  generated by the OPM is not necessarily transitive, and this is true for the  $\sigma - \delta$  specialisation too. For a cycle to occur it is sufficient that there are three alternatives  $(x, t_x)$ ,  $(y, t_y)$  and  $(z, t_z)$  with  $x > z > y$  such that

$$\begin{aligned} y\delta^{t_y} &> x\delta^{t_x} + \sigma \\ z\delta^{t_z} &\leq x\delta^{t_x} + \sigma \\ x\delta^{t_x} &\leq z\delta^{t_z} + \sigma \\ y\delta^{t_y} &\leq z\delta^{t_z} + \sigma \\ z\delta^{t_z} &\leq y\delta^{t_y} + \sigma \end{aligned}$$

so that  $(y, t_y) \succ^* (x, t_x)$  but  $(x, t_x) \sim (z, t_z)$  and  $(y, t_y) \sim (z, t_z)$  by the Primary Criterion. By the Secondary Criterion  $(x, t_x) \succ^* (z, t_z)$  and  $(z, t_z) \succ^* (y, t_y)$ .

The inequalities above are possible whenever

$$y\delta^{t_y} > \sigma + x\delta^{t_x} \geq z\delta^{t_z}$$

## 4.2 Preference reversal

We now study under what conditions the  $\sigma - \delta$  model generates preference reversals in a discrete model. Recall that preference reversal at time  $t'$  is the shorthand for the following situation

$$\begin{aligned} (x, t) &\succ^* (y, t') \\ (y, t' + t'') &\succ^* (x, t + t'') \end{aligned}$$

---

<sup>10</sup>Similarly, our friend Faye likes a sweet at the end of a meal but nonetheless opts for a synthetic sweetener with her coffee.

where  $t' \geq t$ . Then if preferences can be  $u - \sigma$  represented, conditions for preference reversal as expressed above are

$$\begin{aligned} u(x, t) &> u(y, t') + \sigma(y, t') \\ u(x, t + t'') &\leq u(y, t' + t'') + \sigma(y, t' + t'') \end{aligned} \quad (1)$$

Turning to the case of exponential discounting and constant  $\sigma$ , let  $u(x, t) = \delta^t x$ , with  $\delta \in (0, 1)$ . System 1 above simplifies to

$$\begin{aligned} x\delta^t &> y\delta^{t'} + \sigma \\ x\delta^{t+t''} &\leq y\delta^{t'+t''} + \sigma \end{aligned}$$

or:

$$\sigma \in \left[ (x\delta^t - y\delta^{t'}) \delta^{t''}, x\delta^t - y\delta^{t'} \right] \equiv I(x, y, \delta, t, t', t'') \quad (2)$$

So the vagueness factor  $\sigma$  must be sufficiently small to make the smaller amount initially more attractive than the larger, later one; and sufficiently large so that the same comparison is resolved in the opposite direction when all alternatives are pushed forward by  $t''$ . When  $\sigma$  lies in the interval  $I(x, y, \delta, t, t', t'')$  the present discounted value criterion is not sufficiently precise to rank  $(x, t + t'')$  and  $(y, t' + t'')$ . The individual resorts to the Secondary Criterion, and prefers  $(y, t' + t'')$ .

Given  $\sigma$  and the preference at the initial date, preference reversal will eventually occur after a sufficiently long time. This is due to the fact that although  $\sigma$  is constant in absolute terms, the vagueness *relative* to the utilities involved is increasing with time, as the discounted values of the two amounts (and so their difference) tends to 0 as time tends to infinity. Note that in our setting preference reversal is possible whatever the date for the earliest alternative, whereas the  $\beta - \delta$  model can explain preference reversal only if the earliest alternative occurs at time 0. Preference reversal may persist even when  $\delta$  tends to one.

Finally, although in our model the primary ordering is not assumed to be a ‘time preference’ in the technical sense of Masatlioglu and Ok [17], our justification of preference reversal is *not* due to this feature: preference reversal can occur even when within each period primary preferences are complete, transitive and unchanging through time, and satisfy additional continuity conditions as required in that paper.

### 4.3 The ‘now or never’ fallacy

In Giovanni Verga’s [37] short story ‘La roba’ (‘Property’), the peasant Mazzaró leads a miserable life, supporting himself just on bread and onions while at the same time

accumulating a vast wealth by working hard and abstaining from any ‘unnecessary’ consumption. Approaching the end of his life, the realization that he cannot take ‘the stuff’ with him in afterlife drives him insane. We maintain that the ‘now or never’ argument on which we base many consumption decisions is sometimes a fallacy motivated by the fear of ending in a ‘Mazzaró trap’, even to the point of overdoing it, and regretting the decision afterwards.

It is well known that when preferences fail to satisfy time-consistency they may produce patterns of choice which are normatively disturbing though very familiar. The example of Ulysses and the sirens is the most famous one in the literature on time-inconsistency, originating with Strotz [30]: people with weaker foresight than the epic hero will make intertemporal choice which they will later regret. Here we highlight a different type of lack of foresight and ensuing regret.

The outcome produced by some acts depends on the time at which the act takes place, and in particular the later the timing, the better the outcome.<sup>11</sup> An obvious economic example is the choice when to dissave an accumulating capital. Similarly, for consumption goods think of any high-tech product, whose intrinsic qualities improve with technological progress. The feature common to these situations is that there is a tension between the improvement in the outcome (which gives an incentive to postpone the act), and the suboptimality of delaying the act forever. A similar tension may occur even in situations where the improvement of the outcome with the passage of time is not unbounded, but there is a limited planning horizon. For example, there is a fairly precise moment in time when a good cheese reaches the right level of maturity, or a vintage wine the fullness of its flavour, but the decision maker may wish in any case to enjoy it before that moment, so that over the planning horizon the quality of the good is increasing over time. There will be some optimal timing at which this tension is resolved. Consequently, regret would ensue for a decision maker who did take the relevant action (dissaving, purchasing the computer, uncorking the bottle) before such optimal time. We show below how our  $\sigma - \delta$  model can generate this type of time inconsistency for ‘naive’ decision makers who fail to take into account fully their future preferences.<sup>12</sup>

---

<sup>11</sup>Recall that by ‘outcome’ we mean an element of  $X$ , that is we are *not* referring to a date-outcome pair.

<sup>12</sup>The impact of time inconsistent preferences on behaviour has been analysed in different contexts by O’Donoghue and Rabin ([15] and [16]) and Carrillo and Mariotti [5]. The first two papers focus on procrastination. Specifically [15] deals with the issue of the timing of a given task, while [16] considers the additional problem of choosing which task to undertake, and when. [5] analyse a class of problems in the same vein, where an agent’s consumption produces externalities on his future welfare, and the magnitude of the externality is uncertain. Gul and Pesendorfer [10] explain phenomena such as temptation and self-control without invoking the non-stationarity of preferences.

In these cases preferences are such that agents would gladly forego some current utility (delaying consumption/dissaving) in order to have a future higher enjoyment, but are afraid that in the future they will simply repeat this behaviour, therefore delaying enjoyment to an undesirable extent (from the point of view of the present). This tricks them into a ‘now or never’ fallacy: what is the point of accumulating money just to end up like Mazzaró?

We construct an example where some agents may be naive while others are more sophisticated. While a naive person acts on the basis of the initial preferences and opts for the ‘now’, a sophisticated person correctly anticipates his future ability to consume/dissave at the optimal time, thus avoiding any regret.

The now or never fallacy is, in a sense, the mirror image of the standard ‘Ulysses and the sirens’ kind of self-control problem. In that case, naive people lack the foresight to anticipate their future inability to self-control, and engage in activities which they later regret. In the now or never fallacy, conversely, naive people lack the foresight to anticipate their future *ability* to self-control, and refrain from activities which ex-post they would rather have engaged in.

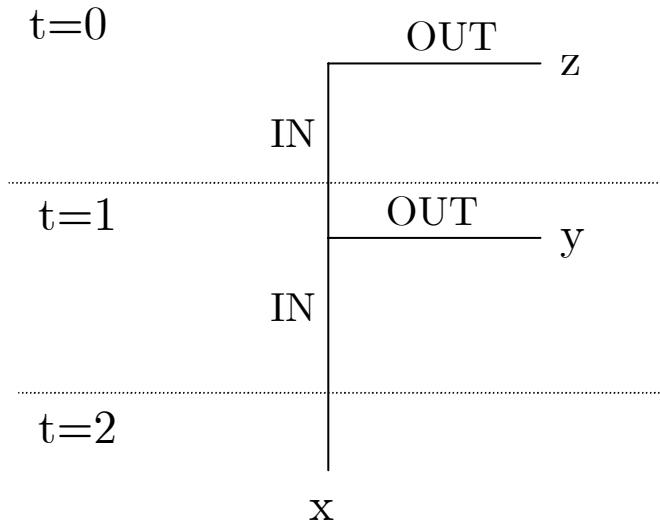


Figure 2: The ‘now or never fallacy’ decision problem, with  $x > y > z$ .

Formally, the intuition of the examples above can be roughly captured by the description of figure 2, where  $x > y > z$ . By choosing *IN* at the outset, the agent foregoes an immediate gratification  $z$  in favour of some higher future outcome (either  $x$  or  $y$ ). In the next period, choosing *OUT* ends the decision problem (uncork the bottle), whereas choosing *IN* again implies a delayed but even higher utility in the future.

Now suppose that for this agent preferences are such that the following holds:

- $z > x\delta^2 + \sigma$ ;
- $x\delta^2 \leq y\delta + \sigma$  and  $y\delta \leq x\delta^2 + \sigma$ ;
- $y\delta \leq z + \sigma$  and  $z \leq y\delta + \sigma$ ;
- $y > x\delta + \sigma$ .

We assume that our naive agents are still sophisticated enough that - given the preferences they are using - they reason consistently in the sense of subgame perfection: they behave optimally in the current period given the anticipated optimal behavior in the prosecution of the decision problem *as they perceive it now*<sup>13</sup>. Then the above inequalities imply that, in period 0, *OUT* at time 0 is preferred outright to *IN* all the way (first condition); *IN* all the way is preferred to *IN* at zero followed by *OUT* in period 1 by the Secondary criterion (second condition); *IN* at zero followed by *OUT* in period 1 is preferred to *OUT* immediately by the Secondary criterion (third condition). Finally the fourth inequality implies that in period 1 the agent prefers *OUT* to *IN*. A naive person will simply assume that his future self will behave according to his current preferences, thus ignoring the last condition. A sophisticated person is instead one who anticipates his preference at 1 for *OUT* over *IN*. As a consequence, a naive person will practice *immediate* restraint by choosing *OUT* at the outset, whereas a sophisticated agent will exercise *delayed* restraint by first choosing *IN* and then *OUT* in period 1.

The ‘bite’ in discriminating between the behaviour of naive and sophisticated agents originates from two facts: first, the initial ordering between  $(x, t)$  and  $(y, t - 1)$  is reversed further down in time (i.e. preferences are non-stationary); second, the initial preferences exhibit non transitivities:  $(z, 0) \succ^* (x, 2) \succ^* (y, 1) \succ^* (z, 0)$ .

These two factors together generate with behavioural patterns which cannot be accounted for by other present biased representations of preferences. To see this latter point, contrast the example above with the prediction the  $\beta - \delta$  model. The initial preferences would be translated in the following requirements:

$$\begin{aligned} z &> \beta\delta^2 x \\ \beta\delta y &> z \\ \beta\delta^2 x &> \beta\delta y \Leftrightarrow \delta x > y \end{aligned}$$

---

<sup>13</sup>An alternative way of describing the decision tree is to consider the corresponding time dependent set of alternatives. Initially (at time 0) the set of alternatives that can be ‘seen’ is  $A_0 = \{(z, 0), (y, 1), (x, 2)\}$ . Conditional upon  $(z, 0)$  being chosen, the problem is over. If instead the action leading to  $(y, 1)$  is selected, then the set of alternatives which can be seen at time 1 is  $A_1 = \{(y, 1), (x, 2)\}$ . A naive person just does not ponder about  $A_1$ .

Note that the first two conditions imply  $\beta\delta y > \beta\delta^2x \Leftrightarrow y > \delta x$ , which contradicts the third one, so that regardless of whether an agent is naive or sophisticated, a person with  $\beta - \delta$  preferences could never find himself confronted with the kind of problems that a  $\sigma - \delta$  agent might face. While the  $\beta - \delta$  model incorporates non-stationarity and thus can easily explain preference reversal, it cannot generate cycles.

We are not claiming that the now or never fallacy is a universal feature of human choices. However, we claim that it is introspectively a possible situation. Just as there are people not as foresighted as Ulysses and engage in an activity wrongly believing they will be able to restrain themselves, so it is plausible that there are naive individuals who refrain from an activity wrongly believing they will *not* be able to restrain themselves. This type of naivete may also be important in explaining low saving rates as well as impulse spending.

#### 4.4 Interpretation of $\delta$ as a SPot ratio

We stated before that  $\delta$  measures the way the individual trades off outcomes across time. This can be made formally more precise in an idealised version of the model in which the space of outcomes is a continuum. For any alternative  $(x, t_x)$  we are not able to define to define a ‘present value’ at time  $t < t_x$  in the standard way, as an alternative at time  $t$  which is indifferent to the later alternative  $(x, t_x)$ . This is because in our theory *no* two alternatives are indifferent, as  $\succ^*$  is a strict order. However, we can define a closely related concept, the Smallest Preferred Outcome at time  $t$ , or *SPot*. This is the outcome  $s(x, t_x, t)$  which solves

$$\delta^t s(x, t_x, t) = \delta^{t_x} x + \sigma$$

Note that  $s(x, t_x, t)$  is not exactly a present value in the sense of being indifferent to  $(x, t_x)$ . However,  $s(x, t_x, t)$  is the infimum of the outcomes available at time  $t$  which are preferred to  $x$ . To see this, assume that  $\sigma$  is sufficiently small that  $s(x, t_x, t) < x$ . Then all outcomes at time  $t$  which are greater than  $s(x, t_x, t)$  are preferred to  $(x, t_x)$  by the Primary Criterion, and  $(x, t_x)$  is preferred to all outcomes at time  $t$  which are not greater than  $s(x, t_x, t)$  by either the Primary or the Secondary Criterion. So a SPot expresses an individual’s ‘willingness to pay’ to anticipate an alternative to time  $t$ . Then for  $0 \leq t-1, t < t_x$  we have  $\delta^t s(x, t_x, t) = \delta^{t_x} x + \sigma = \delta^{t-1} s(x, t_x, t-1)$ , so that  $\delta$  can be written as a ratio of SPot’s:

$$\delta = \frac{s(x, t_x, t-1)}{s(x, t_x, t)}$$

Therefore  $\delta$  can be interpreted as the (subjective) price to anticipate a given alternative  $(x, t_x)$  by one additional period from  $t+1$  to  $t$ .

An analogous definition can be given if  $s(x, t_x, t) > x$ .

These comments on interpretation are also useful for an *empirical estimation* of the parameters  $\delta$  and  $\sigma$ . What the above shows is that they can be elicited by asking subjects questions about willingness to pay in a way similar to the standard procedures used to elicit discount factors.

## 5 Discussion and concluding remarks

Recent evidence has posed a challenge to the traditional model of exponential discounting used to formalise time preferences. With the few notable exceptions discussed earlier, economists have mainly faced this challenge by modifying the exponential discounting function to a hyperbolic one. We have proposed that, instead, the evidence calls for a framework that is adherent to the cognitive procedures human decision makers adopt when making choices involving time. We have highlighted one particular cognitive shortcoming in evaluating future events (vagueness) and suggested one decision heuristics to resolve situations of vagueness. Even focussing on the simplest possible model in the class of our representations, and the one with the smallest departure from exponential discounting, we can explain several apparent paradoxes of decision making. Our modelling approach suggests in a sense a compromise between an ‘economic’ view based on rational orderings, and a ‘psychological’ view, based on heuristics.

Our theory is different from the standard one not so much because of a functional form change, but because it does not rely on the maximisation of a transitive ordering. In the paper we have mainly focussed on the  $\sigma - \delta$  specialisation of the theory because it is the simplest departure from the standard exponential discounting model which is still able to accommodate a host of anomalies. Clearly, however, *any* two-parameter model is valid only in a restricted context<sup>14</sup>. For instance, the  $\sigma - \delta$  model cannot account for ‘size effects’ of the following type: suppose that  $\delta = .7$  and  $\sigma = 1$ , so that  $(9, 0)$  is preferred to  $(10, 1)$  but  $(10, 3)$  is preferred to  $(9, 2)$ , so that preference reversal occurs<sup>15</sup>. Suppose now we double the stakes, and thus look at comparisons between  $(18, 0)$  and  $(20, 1)$ , and between  $(18, 2)$  and  $(20, 3)$ . Now it is easy to verify that the  $\sigma - \delta$  models predicts a preference for the smaller, earlier amount in both comparisons, and preference reversal no longer occurs<sup>16</sup>. To some this might appear contrary to introspection: if for small

---

<sup>14</sup>The same applies to the  $\beta - \delta$  specialisation of the the hyperbolic model, which for instance can account for preference reversal only when the initial time 0 is involved.

<sup>15</sup>Note that  $u(9, 0) = 9 > (.7) 10 + 1 = 8 = u(10, 1) + \sigma$ , while  $u(9, 2) = 9 (.7)^2 = 4.41 < 10 (.7)^3 + 1 = 4.43 = u(10, 3) + \sigma$ .

<sup>16</sup>Now  $u(18, 0) = 18 > (.7) 20 + 1 = 15 = u(20, 1) + \sigma$ , and also  $u(18, 2) = 18 (.7)^2 = 8.82 >$

stakes one is willing to wait for a future monetary gain (so that  $(10, 3)$  is preferred to  $(9, 2)$ ), then he should be even more willing if this gain is doubled. First of all one could argue that in the above example the misprediction is due to a misspecification of the taste parameters  $\delta$  and  $\sigma$ : preference reversal exhibited in both comparisons would for instance be compatible<sup>17</sup> with  $\sigma \in [1.96, 2]$ . More importantly, the apparent misprediction of the example is based on the implicit assumption that *differences* matter, and that some sort of internal consistency based on similarity is relied upon, so that if a unitary gain calls for preference reversal, so must a gain which is double in size<sup>18</sup>. Still, there is nothing particularly unreasonable in finding it more difficult to make comparisons when high stakes are involved. So for instance the vagueness term  $\sigma(y, t)$  can be modelled as a *proportion* of the outcome, that is  $\sigma(y, t) = \sigma y$  for some percentage  $\sigma$ . Then we have that  $x > \delta^t y + \sigma y$  if and only if  $\lambda x > \delta^t (\lambda y) + \sigma (\lambda y)$  for any positive scaling factor  $\lambda$ , so that the ‘misprediction’ of the numerical example above disappears<sup>19</sup>.

Note that this specialisation belongs to our general class of representations. The crucial aspect is that the vagueness term enters *additively*. Therefore *any* form of discounting in the utility function  $u$  is compatible with increasing relative vagueness whenever the vagueness term  $\sigma(y, t)$  is time-independent (or even when, if utility is discounted exponentially,  $\sigma(y, t)$  increases less than exponentially with time). It is not necessary that  $\sigma(y, t)$  is a constant, which is instead our assumption in the  $\sigma - \delta$  model.

To sum up, our theory is more general than the  $\sigma - \delta$  specialisation. The exponential discounting part of it is not central - indeed, it may well be that hyperbolic discounting is the ‘psychologically correct’ way of dealing with preference for anticipation. In this light, our core ideas may be viewed as complementary, rather than in contrast, to the class of hyperbolic discounting models. We believe they bring to the fore some fundamental cognitive aspects of choice over time that the existing theories do not consider.

Our model is limited in that so far it does not deal with utility streams. Although the existing models of discounting have been readily applied to streams, we believe that this topic requires some special considerations. An important contribution in the same spirit as ours is that by Jehiel and Lilico [11]. They model the decision maker’s bounded rationality as limited foresight, so in this sense their agents are also ‘vague’ about the future. We

---


$$20(.7)^3 + 1 = 7.86 = u(20, 3) + \sigma.$$

<sup>17</sup>Now the required inequalities are  $9 - (.7)10 = 2 > \sigma$  and  $9(.7)^2 - 10(.7)^3 = .98 \leq \sigma$  for the ‘small stakes’ comparison; and  $18 - (.7)20 = 4 > \sigma$  and  $18(.7)^2 - 20(.7)^3 = 1.96 \leq \sigma$  for the ‘high stakes’ comparison.

<sup>18</sup>Incidentally, this also seem to rely on a ‘similarity’ heuristic based on differences in one component - our initial premise was the rejection of this assumption *as a starting point*.

<sup>19</sup>Preference reversal (between times 0 and  $k$ ) is obtained when  $y > x$  and  $x > \delta^t y + \sigma y$  holds simultaneously with  $\delta^k x \leq \delta^{t+k} y + \sigma y$ , or  $\sigma + \delta^t < \frac{x}{y} \leq \frac{\sigma}{\delta^k} + \delta^t$ .

leave the extension of our own framework to utility streams to future research.

## 6 Appendix

**Proposition 3** Let  $X = T = [0, 1]$ . Let Discrimination, Time Monotonicity (i), Outcome Monotonicity (i), Continuity and Betweenness hold. Then  $\succ$  is  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and  $u$  non-decreasing in outcome and non-increasing in time.

**Proof.** Define  $u(a) = \lambda(L(a))$ , where  $\lambda$  denotes the Lebesgue measure. By Time Monotonicity (i) and Outcome Monotonicity (i), the function  $u$  is weakly monotonic in the desired direction with respect to time and outcomes. Let

$$L^*(a) = \lim_{b \in V(a)} L(b)$$

This limit is well defined since by Discrimination the sets  $L(b)$  in the formula are nested. Denote  $\lambda^*(a) = \lambda(L^*(a))$ , and define

$$\sigma(a) = \max\{0, \lambda^*(a) - u(a)\}$$

We show that  $u$  and  $\sigma$  as defined above represent  $\succ$ . Suppose  $a \succ b$ . We have  $L(b) \subset L(a)$  by the transitivity and irreflexivity of  $\succ$ , and by Betweenness  $\lambda(L(a)) > \lambda(L(b))$ , that is  $u(a) > u(b)$ . Now suppose in negation that

$$u(a) \leq u(b) + \sigma(b)$$

By construction  $\sigma(b) = \max\{0, \lambda^*(b) - u(b)\}$ . If  $\sigma(b) = 0$  we get an immediate contradiction with  $u(a) > u(b)$ . So let instead  $\sigma(b) = \lambda^*(b) - u(b)$  and therefore

$$u(a) \leq u(b) + \sigma(b) = \lambda^*(b)$$

Suppose there exists  $b^* \in V(b)$  such that  $\lambda(L(b^*)) = \lambda^*(b)$ . Then we have  $u(a) \leq u(b^*)$ . By Discrimination  $L(a) \subseteq L(b^*)$ , so that  $b \in L(a)$  contradicts  $b \in V(b^*)$ . We show finally that such a  $b^*$  exists. If not, then  $L^*(b) \notin \{L(b') | b' \in V(b)\}$  and there must exist a sequence  $\{b_n\}$ , with  $b_n \in V(b)$  for all  $n$ , and with  $\lim b_n = b^* \notin V(b)$ . If  $b^* \succ b$  then (i) of Continuity is violated, and if  $b \succ b^*$  then (ii) of Continuity is violated.

The proof concludes as the proof of proposition 1. ■

**Proposition 4** Let  $X$  and  $T$  be countable sets. Let Discrimination, Time Monotonicity (i), Outcome Monotonicity (i) hold. Suppose that  $(\lim \{a_n\}) \in V(a)$  for any sequence

$\{a_n\}$  with  $a_n \in V(a)$ . Then  $\succ$  is  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and  $u$  non-decreasing in outcome and non-increasing in time.

**Proof.** The proof is the same as the proof of proposition 3, where  $\lambda(a)$  defines a numerical representation of the partial order of strict set inclusion instead of the Lebesgue measure (see e.g. Fishburn [6] for the existence of such a numerical representation). Note that the limit property in the statement implies the existence of a  $b^* \in V(b)$  such that  $\lambda(L(b^*)) = \lambda^*(b)$ . ■

**Proposition 5** Let  $\succ$  be  $(u, \sigma)$ -representable with  $\sigma(a) \geq 0$  for all  $a \in A$  and let Time Monotonicity (i) and (ii), and Outcome Monotonicity (i) and (ii) hold. If  $\sigma$  can be chosen to be non-increasing in outcome, then  $u$  can be chosen to be non-decreasing in outcome at an arbitrarily large number of alternatives, and if  $\sigma$  can be chosen to be non-decreasing in time then  $u$  can be chosen to be non-increasing in time at an arbitrarily large number of alternatives.

**Proof.** Fix a representation  $u$  and  $\sigma$ , and suppose that  $u$  is decreasing in outcome at some point, that is there exist  $x_2 P x_1$  and  $t$  such that  $u(x_1, t) > u(x_2, t)$ . Consider the transformation  $v$  of  $u$  given by

$$v(x, t') = \begin{cases} u(x_2, t) & \text{if } x = x_1 \text{ and } t' = t \\ u(x, t') & \text{otherwise} \end{cases}$$

This represents the same preferences as  $u$ . To see this, note first that, for any  $(y, t') \succ (x_1, t)$ , since  $u$  and  $\sigma$  represent  $\succ$  we have  $u(y, t') > u(x_1, t) + \sigma(x_1, t)$ , and therefore

$$v(y, t') = u(y, t') > u(x_1, t) + \sigma(x_1, t) > u(x_2, t) + \sigma(x_1, t) = v(x_1, t) + \sigma(x_1, t)$$

so that  $v$  and  $\sigma$  still represent  $\succ$  for all alternatives which are preferred to  $(x_1, t)$ .

Next consider  $(x_1, t) \succ (y, t')$ . Suppose by contradiction that  $v(x_1, t) \leq u(y, t') + \sigma(y, t')$ . Then by construction  $u(x_2, t) \leq u(y, t') + \sigma(y, t')$ , so that  $(x_2, t) \not\succ (y, t')$ . This together with  $(x_1, t) \succ (y, t')$  contradicts part (i) of Outcome Monotonicity.

Consider now  $(x_1, t) \sim (y, t')$ . Suppose by contradiction that  $u(y, t') > v(x_1, t) + \sigma(x_1, t)$ . Then by construction  $u(y, t') > u(x_2, t) + \sigma(x_1, t) \geq u(x_2, t) + \sigma(x_2, t)$ , where the second inequality follows from the monotonicity assumption on  $\sigma$  in the statement. Then  $(y, t') \succ (x_2, t)$ , which together with  $(x_1, t) \sim (y, t')$  contradicts part (ii) of Outcome Monotonicity.

This procedure can be repeated for an arbitrarily large number of pairs of alternatives where non-monotonicity occurs.

Finally, to prove that if  $\sigma$  can be chosen to be non-decreasing in time then  $u$  can be chosen to be non-increasing in time arguments essentially identical to the ones above

apply by using Time Monotonicity in place of Outcome Monotonicity and the following transformation

$$v(x', t) = \begin{cases} u(x, t_1) & \text{if } x' = x \text{ and } t = t_2 \\ u(x', t) & \text{otherwise} \end{cases}$$

at all points such that  $u(x, t_2) > u(x, t_1)$  with  $t_2 > t_1$ . ■

## References

- [1] Aizpurúa, J. M., T. Ichishi, J. Nieto and J. R. Uriarte, 1993 “Similarity and preferences in the space of simple lotteries”, *Journal of Risk and Uncertainty*, 6: 289-297.
- [2] Aizpurúa, J. M., J. Nieto and J. R. Uriarte, 1990, “Choice procedure consistent with similarity relations” *Theory and Decisions*, 29: 235-354.
- [3] Ainslie, G., 1975 “Specious reward: a behavioral theory of impulsiveness and impulsive control”, *Psychological Bulletin*, 82: 463-496.
- [4] von Böhm-Bawerk, E., 1889/1970, *Capital and Interest*, Libertarian Press, South Holland.
- [5] Carrillo, J. D. and T. Mariotti, 2000, “Strategic ignorance as a self disciplining device”, *Review of Economic Studies*, 67: 529-544.
- [6] Fishburn, P., 1970, *Utility theory for decision making*, Wiley, New York.
- [7] Fishburn, P., and A. Rubinstein, 1982, “Time Preference”, *International Economic Review*, 23: 677-694.
- [8] Frederick, S., G. Loewenstein and T. O’Donoghue, 2002, “Time discounting: a critical review”, mimeo.
- [9] Gilboa, I. and R. Lapson, 1995, “Aggregation of semiorders: intransitive indifference makes a difference”, *Economic Theory*, 5: 109-126.
- [10] Gul, F. and W. Pesendorfer, 2001, “Temptation and self-control”, *Econometrica*, 69:1403-1435.
- [11] Jehiel, P. and A. Lilico, 2002 “Smoking today and stopping tomorrow: a limited foresight perspective”, mimeo, UCL.

- [12] Laibson, D., 1997, “Golden eggs and hyperbolic discounting”, *Quarterly Journal of Economics*, 112: 443-477.
- [13] Loewenstein, G. and D. Prelec, 1992, “Anomalies in intertemporal choice: evidence and interpretation”, *Quarterly Journal of Economics*, 107: p. 573-597.
- [14] Luce, D., 1956 “Semiorders and a theory of utility discrimination”, *Econometrica*, 24:178-191.
- [15] O’Donoghue, T. and M. Rabin, 1999, “Doing it now or later”, *American Economic Review*, 89: 103-124.
- [16] O’Donoghue, T. and M. Rabin, 2001, “Choice and procrastination”, *Quarterly Journal of Economics*, 116:121-160.
- [17] Ok, E. and Y. Masatlioglu, 2003 “A General Theory of Time Preferences”, mimeo, New York University.
- [18] Phelps, E. and R. Pollack, 1968, “On second best national savings and game-equilibrium growth”, *Review of Economic Studies*, 35:201-208.
- [19] Pigou, A.C. , 1920, *The Economics of Welfare*, London, McMillan.
- [20] Pittano, G., 1992, “*Frase fatta capo ha*”: dizionario dei modi di dire, proverbi e locuzioni, Zanichelli, Bologna.
- [21] Read, D. , 2001 “Is Time-Discounting Hyperbolic or Subadditive?”, *Journal of Risk and Uncertainty*, 23:5-32.
- [22] Roelofsma, P. H. and D. Read, 2000 “Intransitive Intertemporal Choice”, *Journal of Behavioral Decision Making*, 13: 161-177.
- [23] Rovira, L. I., 1984, *Spanish proverbs : a survey of Spanish culture and civilization*, University Press of America.
- [24] Rubinstein, A., 1988, “Similarity and Decision Making Under Risk (Is there a utility theory resolution to the Allais Paradox?)”, *Journal of Economic Theory*, 46: 145-153.
- [25] Rubinstein, A., 2000, “Is it ‘Economics and Psychology’? The case of hyperbolic discounting”, mimeo.
- [26] Rubinstein, A., 2001, “A theorist’s view of experiments”, *European Economic Review*, 45: 615-628.

- [27] Shafir, E., I. Simonson and A. Tversky, 1993, “Reason-based choice”, *Cognition*, 49:11-36.
- [28] Simpson, J. and J. Speake, 1992, *Concise Oxford Dictionary of Proverbs*, Oxford University Press.
- [29] Slovic, P., 1975, “Choice between equally valued alternatives”, *Journal of Experimental Psychology: Human Perception and Performance*, 1: 280-287.
- [30] Strotz, R. H., 1956, “Myopia and inconsistency in dynamic utility maximization”, *Review of Economic Studies* 23: 165-180.
- [31] Taylor, A., 1931, *The proverb*, Harvard University Press
- [32] Thaler, R. and H. Shefrin, 1981, “An economic theory of self-control”, *Journal of Political Economy*, 89: 392-410.
- [33] Tversky, A., 1969, “Intransitivity of Preferences”, *Psychological Review*, 76:31-48.
- [34] Tversky; A., P. Slovic and D. Kahneman, 1990, “The Causes of Preference Reversal”, *The American Economic Review*, Vol. 80: 204-217
- [35] Tversky, A., S. Sattath and P. Slovic, 2000, “Contingent weighting in judgement and choice”, in D. Kahneman and A. Tversky (eds.), *Choices, Values and Frames*, Cambridge University Press, Cambridge, UK.
- [36] Uriarte, J. R., 1999 “Decision-making under risk: editing procedures based on correlated similarities, and preference overdetermination” *Review of Economic Design*, 4: 1-12.
- [37] Verga, G. (1883) “La roba”, in *Novelle rusticane*, 1883 (translation from the Italian by D.H. Lawrence: “Property”, in *Short Sicilian novels*, Dedalus, 1984, London).
- [38] Vilà, X., 1998, “On the Intransitivity of Preferences Consistent with Similarity Relations”, *Journal of Economic Theory*, 79: 281-287.