# From Decision in Risk to Decision in Time - and Return: A Restatement of Probability Discounting

Marc-Arthur Diaye, André Lapidus, Christian Schmidt

# From Decision in Risk to Decision in Time - and Return: A Restatement of Probability Discounting

Marc-Arthur Diaye*, André Lapidus** and Christian Schmidt***

June 01, 2021

**Abstract**

This paper aims at restating, in a decision theory framework, the results of some significant contributions of the literature on probability discounting that followed the publication of the pioneering article by Rachlin *et al.* (1991). We provide a restatement of probability discounting in terms of rank-dependent utility, in which the utilities of the outcomes of $n$-issues lotteries are weighted by probabilities transformed after their transposition into time-delays. This formalism makes the typical cases of rationality in time and in risk mutually exclusive, but allows looser types of rationality. The resulting attitude toward probability and toward risk are then determined in relation to the values of the two parameters involved in the procedure of probability discounting.

*Keywords*: Probability discounting, time discounting, logarithmic time perception, rank-dependent utility, rationality, attitude toward probabilities, attitude toward risk.

*JEL classification*: D81, D91, D15.

## 1 Introduction

The existence of significant parallels between decision in time and decision in risk is rather intuitive because of the formal similarities between standard discounted and expected utility. But the more specific thesis that delayed reward and probable reward could be treated in the same way because, contrary to a common view, they refer to the same matter, is less familiar. It seems to have been first explored by psychologists like J.B. Rotter (1954) for whom delays of gratification could be regarded as involving risky rewards by their very nature. Later, some authors like D. Prelec and G.F. Loewenstein (1991), initiated a large stream of works by arguing, on the basis of anomalies observed in both expected utility and discounted utility models, that a delayed reward and a probable reward could be dealt with in the same way, within a multi-attribute choice model. In the same time, H. Rachlin and his co-authors (H. Rachlin, A. Raineri and D. Cross 1991, in the continuation of H. Rachlin, A.W. Logue, J. Gibbon, M. Frankel 1986) developed, in a seminal paper which accounts for experiments with college undergraduates, the idea that a probable reward could be viewed as a delayed reward[1], discounted to obtain its present value, provided probabilities, regarded as "odds-against", are transposed into delays. This approach became widely spread (see, for instance, among others, H. Rachlin and E. Siegel 1994, H. Rachlin, E. Siegel and D. Cross, 1994, P. Ostaszewski, L. Green and J. Myerson 1998, H. Rachlin, J. Brown and D. Cross 2000, L. Green and J. Myerson 2004, T. Takahashi 2005, R. Yi, X. de la Piedad, W.K. Bickel 2006) and gave rise to what was first called "probabilistic discounting" by Rachlin *et al.* (1991).

*Sorbonne Center for Economics (CES), University Paris 1 Panthéon-Sorbonne – 106-112, boulevard de l'Hôpital – 75647 Paris Cedex 13 – France. E-mail: marc-arthur.diaye@univ-paris1.fr.

**PHARE, University Paris 1 Panthéon-Sorbonne – 106-112, boulevard de l'Hôpital – 75647 Paris Cedex 13 – France. E-mail: andre.lapidus@univ-paris1.fr

***PHARE and University Paris-Dauphine – Place de Lattre de Tassigny – 75775 Paris cedex 16 – France. E-mail: schmidtchristian.f@orange.fr

[1] Such relation between probability and delay was already formally in use as early as 1713 in what we know as "Bernoulli trials", named after Jacob Bernoulli in *Ars conjectandi*.

The discounting function which aimed to account for decision under risk was assumed to be of a hyperbolic kind[2] on the basis of arguments either empirical, or pertaining to the shape of the relation between the reward and the rate of reward. From an analytical viewpoint, something new occured with the publication of a paper by D.O. Cajueiro (2006) who first introduced a hyperbolic discounting function based in the deformed algebra inspired by Tsallis' non-extensive thermodynamics (C. Tsallis 1994), the $q$-exponential function[3], specially relevant to account for increasing impatience. In the continuation of Cajueiro (2006), T. Takahashi, either alone (Takahashi 2007a, 2011) or with various authors (Takahashi *et al.* 2012, 2013) took over the $q$-exponential function to account for time discounting as well as probability discounting. Meanwhile, the same authors focused on the nature of the delay associated to probabilities in probability discounting, focusing on the distinction between physical and perceived waiting time. A classical approach, regarding the way an external stimulus is scaled into an internal representation of sensation, which was initiated by Weber and Fechner in the second half of the 19th century in psychophysics, concluded that the relation was logarithmic. Nearly a century later, the issue was revived by S. Stevens (1957), who discussed the possibility of an alternative (power functions) to the logarithmic relation. More recently, some authors (see S. Dehaene 2003) have given a neural basis to the view that our mental scaling be logarithmic. In line with this perspective, Takahashi and his co-authors supported the view that the perceived waiting time was logarithmically related to the physical waiting time (Takahashi 2005, 2011; Takahashi *et al.* 2012). In Takahashi (2005) and in Takahashi *et al.* (2012), the reward was submited to an exponential discount, but relatively not to *physical* waiting time, but to *perceived* waiting time. Relatively to physical waiting time, this resulted in a general hyperbolic discounting function in Takahashi (2005) transposed into a $q$-exponential discounting function in Takahashi *et al.* (2012). It is obvious that, as a consequence, the outcome of the operation was, like in D. Kahneman and A. Tversky (1979) and with similar consequences, a transformation of the decision weight of the probability associated with the reward.

A common point of this literature (with, of course, the notable exception of papers like this of Prelec and Loewenstein (1991)) is that its main concern was to identify a few typical relations consistent with the results of limited experiments related to choices under risk or over time. From this point of view, it can rightly be considered a success story. But on the other hand, the theoretical support of these experimental results is often limited by what is strictly required and presented in a piecemeal way, according to the needs of the experiments. For instance, the idea of a logarithmic perception of physical time appeared as early as 2005 in Takahashi's work, in a paper devoted to time discounting, not to decision in risk. Its integration to a wider representation leading to a probability weighting function of which Prelec (1998) was a special case only occured six years later (Takahashi 2011; see also Takahashi *et al.* 2012). In the same way, the systematic use of 2-issues lotteries in which one looses or wins, is appropriate to dealing with issues like the comparison of the respective effects of exponential and hyperbolic discounting on the discounted value of a reward, or of the distortion of the probability of obtaining a reward induced by probability discounting. However, this limitation to simple 2-isues lotteries has significant consequences regarding the way probabilities are perceived. At last, the issue of the desirability of the reward is not being addressed head-on. References to the pioneering work of Kahneman and Tversky (1979) are quite frequent, but they usually concerns the weighting of probabilities, not the value function that would lead us to consider that our preferences relate not to a state (through a utility function), but to a difference with respect to the *statu quo* (the value function).

In the follow-up of this article, we provide a restatement of probability discounting in which probabilities are transformed into expected delays before winning, but where (i) the usual case of a 2-issues lottery is extended to the more general case of discrete random variables with finite support and (ii) a utility function is explicitly introduced in the analysis, so that we come to a rank-dependent utility approach[4] (Section 2). The resulting formalism makes the typical standard

[2]Rachlin, Raineri and Cross (1991), for instance, explicitly refer to the discounting function $(1+\alpha t)^{-1}$ (with $t$ denoting time and $\alpha$ a discounting parameter) proposed by J.E. Masur (1987). The same function was previously introduced in 1981 by R. Herrnstein. Rachlin, Siegel and Cross (1994) proposed a general hyperbolic discounting function of the type $(1+\alpha t)^{-\beta}$ - which may be thought rather close to the function introduced by Loewenstein and Prelec (1992).

[3]The $q$-exponential function $\exp_q$ is defined as $\exp_q(\alpha t) = (1+(1-q)\,\alpha t)^{-\frac{1}{1-q}}$. See Cajueiro (2006), p. 386.

[4]A value function, measuring the differences with a state of reference, could have been used instead of a utility

cases of rationality in time and in risk mutually exclusive, but allows looser types of rationality, involved in the axiomatisation of generalised hyperbolic discounting and of rank-dependent utility. According to the values of its parameters, we find the classical properties of probability weighting functions, expressing pessimism or optimism with regard to probabilities. In combination with the utility function, probability discounting gives rise to the various types of attitudes toward risk, one of the parameters playing a crucial part as an index of pessimism (Section 3).

# 2 Extending probability discounting

As an approach to decision under risk through a specific valuation of lotteries, the probability discounting approach which emerge from the pioneering work of Rachlin *et al.* (1991) might be viewed as a four steps procedure, involving 1. the transposition of a probability into a physical delay; 2. the transformation of this physical in a perceived delay; 3. the assessment of a resulting temporal discounting; and 4. the transformation of a discounted delayed value into the utility of a probable reward. The four steps of this procedure are outlined below.

## 2.1 *From the probability of gain to a physical delay before this gain: a Bernoulli trial transposition*

The usual framework of probability discounting is, more or less explicitly, this of a representation of decision under risk where the set $\Lambda$ of probability distributions is typically defined over $\{0, x\}$, $x$ being the possible gain, of probability $p$, of a 2-issues lottery $L$ belonging to $\Lambda$. After Rachlin *et al.* (1991), a common feature of the probability discounting contributions is that $L$ is related to the valuation of a decision through a waiting time $l$, which can be interpreted, as already noted by Rachlin *et al.* (1986, p. 36), as "odds against" in repeated gambles. Though rather intuitive, this interpretation could be given a firmer basis than the usual one, which draws on the comparison with a gambler betting on a horse race, in terms of repeated Bernoulli trials. It is well known that the expected value of the random variable representing the number of trials before winning (the winning trial included) is $1/p$. If we take the interval between two trials as the unit of time, the expected value of the physical delay $l$ before the winning trial is therefore given by:

$$l = \frac{1}{p} - 1 = \frac{1-p}{p} \tag{1}$$

Such representation of the link between probability and physical delay has been currently admitted, since Rachlin *et al.* (1991), as the initial moment of a procedure leading to transform probabilities. Insofar as we remain in the framework of 2-issues lotteries, and as the counterpart of the transformation of the probability $p$ of success is a parallel and consistant transformation of the probability $1-p$ of failure, the immediate link in (1) between probability and delay is not contentious. But regarding the more general case of $n$-issues lotteries is less simple. Assume these lotteries $L$ are the laws of probability of discrete random variables $X$ with finite support:

$$L = (x_1, ..., x_i, ..., x_n; p_1, ..., p_i, ..., p_n) \tag{2}$$

in which the outcomes $x_i$ are ranked in increasing order, $x_1 < ... < x_i < ... < x_n$, and $\sum_{i=1}^{n} p_i = 1$.

Let $G$ be the decumulative distribution function of the random variable $X$ whose probability law is given by the lottery $L$: $G(x_i) = \Pr(X \geq x_i)$. It is obvious that $G(x_1) = 1$ and $G(x_n) = p_n$. Consider now not the isolated probability $p_i$ of obtaining $x_i$, but the probability of obtaining *at least* $x_i$, that is, $G(x_i)$. We can derive from $G(x_i)$ a Bernoulli trial whose issues are either sucess, with an outcome between $x_i$ and $x_n$ both included, or failure, with an outcome between $x_1$ and $x_{i-1}$ also included. $G(x_i)$ is therefore the probability of success, and $F(x_i) = 1 - G(x_i)$ the probability of failure ($F(x_i)$ standing for the usual cumulative distribution function). The expected number of Bernoulli trials to obtain one success (that is, getting at least $x_i$) is $1/G(x_i)$. And going on

function. The result would have been a variant of Kahneman and Tversky's cumulative prospect theory. We have preferred the methodologically simpler representation of rank-dependent utility, whose transposition to cumulative prospect theory can be easily performed.

transposing probability into a physical delay before winning in a repeated gamble like in (1), the average delay $l_i$ for success, that is for obtaining *at least* $x_i$ is given by:

$$l_i = \frac{1 - G(x_i)}{G(x_i)} \quad (i = 1, ..., n) \tag{3}$$

## 2.2 *From a physical to a perceived delay: a logarithmic treatment*

As far as $p$ is an objective probability, $l$ can be viewed as "physical waiting time" (Takahashi *et al.* 2012, p. 13). Drawing on the reintroduction of Fechnerian-like perspectives in psychophysics (see S. Dehaene 2003), Takahashi and his co-authors assume that in a 2-issues lottery, the subjectively perceived waiting time $\tau$ is a logarithmic function of the physical waiting time

$$\tau = a \ln(1 + bl) \tag{4}$$

with $a, b > 0$ (see, for instance, Takahashi 2005, p. 692). The same principles hold in the general case of a $n$-issues lottery: the subjectively perceived waiting time $\tau_i$ before winning at least $x_i$ is logarithmically related to the physical waiting time:

$$\tau_i = a \ln(1 + bl_i) \quad (i = 1, ..., n) \tag{5}$$

Replacing the physical delay by decumulated probability (i.e., the probability of winning at least a certain outcome) like in (3), the probability of winning at least $x_i$ is related as follows to the perceived delay before winning at least $x_i$:

$$\tau_i = a \ln\left(1 + b\frac{1 - G(x_i)}{G(x_i)}\right) \quad (i = 1, ..., n) \tag{6}$$

## 2.3 *From perceived time discounting to physical time discounting*

The third step provides a separate treatment of time discounting. In the case of a 2-issues lottery explored by standard litterature on probability discounting, things are rather simple. The basic idea is this of an exponential discounting whose argument is the perceived delay $\tau$, instead of the physical delay $l$:

$$\mu = \exp(-r\tau) \tag{7}$$

where $\mu$ and $r$ stand respectively for the discounting factor and the discount rate[5] for an outcome $x$, whose expected perceived delay before winning it, is $\tau$. From (4) and (7) we therefore have:

$$\mu = (1 + bl)^{-ra} \tag{8}$$

which amounts to a generalized hyperbolic discounting factor, like in Loewenstein and Prelec (1992)[6]. Exponential discounting, relatively to perceived time, has therefore generated hyperbolic discounting, relatively to physical time.

However, such determination of the discounting factor would be seriously flawed if extended as such to $n$-issues lotteries: if, drawing on (5), $\exp(-r\tau_i) = (1 + bl_i)^{-ra}$ can rightly be viewed

[5] A separate discount rate $r$ related to perceived time is generally missing in the usual literature on probability discounting (see, for instance, Takahashi 2005; but Takahashi *et al.* (2012, p. 12) seem to have done a choice similar to ours). This might be explained by the integration of the relevant infomation in the parameter $a$ in the relation between perceived and physical time. The drawbacks of such way of processing is that it does not make any distinction between discounting in time and perceiving time. This is why we have chosen to make the discount rate explicit.

[6] The generalized hyperbolic discounting factor in Loewenstein and Prelec (1992) writes $(1 + \alpha l)^{-\frac{\beta}{\alpha}}$. Setting $b = \alpha$ and $ra = \beta/\alpha$ enables to find the formulation of (8).

as a discounting factor, it depends on the expected time (perceived or physical) before winning *at least* $x_i$ - not before winning *exactly* $x_i$. The discounting factor associated to the outcome $x_i$ is therefore the difference between two discounting factors: the one related to the expected time before winning at least $x_i$ and the one related to the expected time before obtaining strictly more than $x_i$, that is at least $x_{i+1}$. So that, assuming that $l_{n+1} \to +\infty$:

$$\begin{aligned} &\mu_i = (1+bl_i)^{-ra} - (1+bl_{i+1})^{-ra} \\ &\quad (i=1,...,n) \end{aligned} \tag{9}$$

After the work of Cajueiro in 2006, the expression of the discounting factor has been currently rewritten, through a change in the parameters, as a $q$-exponential discounting based on Tsallis' statistics. This change leads to set a pair of alternative parameters, $k$ and $q$ defined as $k = rab$ and $q = 1 - 1/ra$. Extending this redefinition to the expression of $\mu_i$, (9) can be rewritten as[7]:

$$\begin{aligned} &\mu_i = \psi(l_i) - \psi(l_{i+1}) \\ &\text{where } \psi(l_i) = (1 + k(1-q)\,l_i)^{-\frac{1}{1-q}} \\ &(i=1,...,n) \end{aligned} \tag{10}$$

Or, using Cajueiro's notation for $q$-exponential discounting:

$$\begin{aligned} &\mu_i = \exp_q(kl_i) - \exp_q(kl_{i+1}) \\ &(i=1,...,n) \end{aligned} \tag{11}$$

The discounting factor $\mu_i$ can therefore be equivalently expressed as the difference between the values $\psi(l_i)$ and $\psi(l_{i+1})$ of two generalized hyperbolic discountings (10) or, equivalently, between two $q$-exponential discountings (11). Cajueiro's presentation introducing in 2006 $q$-exponential discounting can be found as early as the following year in Takahashi (2007a) and the colleagues with whom he had partnered (see, for instance, Takahashi 2010, Takahashi 2011, Takahashi *et al.* 2012, Takahashi 2013, Takahashi *et al.* 2013). It will be considered that, because of the definition of $a$, $b$ and $r$ in (4) and (5), the parameters $k$ and $q$ are, by construction, such that $k \geq 0$ and $-\infty < q < 1$. The possibility that $q$ is negative does not appear in the article by Cajueiro (2006), nor in that of Takahashi (2007a). However, when he resumes $q$-discounting during the same year or the following year but in an intertemporal choice framework, Takahashi (2007b, 2008) explicitly considers the possibility that $q$ is less than 0[8]. The interpretation of the parameters $k$ and $q$ will be discussed in Section 3.

## 2.4 *From a discounted delayed value to the utility of a probable reward*

The recourse to an explicit representation of the desirability of the reward is lacking in the works on probability discounting cited above. The emphasis placed on the transposition of probabilities into delays, as well as the binary structure of lotteries, justified a minimum treatment allowing to ignore it. It was sufficient to work with a simple function $V(x,t)$ whose two arguments, the outcome $x$ and the delay $t$ before winning had each one only two possible values: $x = 0$ in case of failure or $x = \bar{x}$ in case of success; $t = 0$ for an immediate (because certain) gain, $t = l$ for a delayed reward (because its probability $p$ is such that $l = (1-p)/p$). Assuming that $V(0,t) = 0$, the immediate or certain value of the reward $\bar{x}$ writes $V(\bar{x}, 0)$, and its delayed or with probability $p$, value is $V(\bar{x}, l)$. This is enough to get

$$V(x,l) = \mu V(x,0) \tag{12}$$

which is all we need to focus on the specification and the discussion of the discounting factor $\mu$. But such simplicity must be abandoned when moving on to the more general case of $n$-issues lotteries

[7]Faced with a 2-issues lottery, we find, as a special case, the usual results from the literature on $q$-discounting (see, for instance, Takahashi 2007a) with a discounting factor for the outcome in case of success $\mu = (1 + k(1-q)\,l)^{-\frac{1}{1-q}} = \exp_q(kl)$.

[8]S. Cruz Rambaud and M.J. Munoz Torrecillas (2013) went so far as to propose that $q$ is greater than 1 (see also Munoz Torrecillas *et al.* 2017). Nonetheless, since this would result in the negativity of $r$ or $a$, and the negativity of $b$ if we want to keep $k$ positive, this possibility is excluded in the following of this paper.

which require comparisons between the desirability of the various possible outcomes when they are immediate or certain. This desirability can be represented by an increasing utility function $u$ of $x$, calibrated so that $u(0) = 0$, and defined up to a positive linear transformation. So that the utility of a lottery $U(L)$ can be given, like for utility in time, as the sum of the undiscounted utilities of each possible outcome $u(x_i)$ weighted by its discounting factor $\mu_i$ defined as in (10):

$$U(L) = \sum_{i=1}^{n} \mu_i u(x_i) \tag{13}$$

Now, because of the probability discounting perspective, $\mu_i$ in (13) can be understood either as a discounting factor whose expression is given by $\mu_i = \psi(l_i) - \psi(l_{i+1})$ in (10), or as probability decision weights. Relying on (3) and (10) we get an alternative expression of $\mu_i$, as the decision weight for obtaining an outcome $x_i$. $\mu_i$ is the difference between the transformed probability $G(x_i)$ of winning at least $x_i$ and the transformed probability of winning strictly more than $x_i$, $G(x_{i+1})$:[9]

$$\begin{aligned} &\mu_i = \varphi(G(x_i)) - \varphi(G(x_{i+1})) \\ &\text{where } \varphi(G(x_i)) = \left(1 + k(1-q)\frac{1-G(x_i)}{G(x_i)}\right)^{-\frac{1}{1-q}} \\ &(i = 1, ..., n) \end{aligned} \tag{14}$$

It can be shown that the probability weighting function $\varphi$ is an increasing transformation of $[0, 1]$ into itself with the following properties:

$$\begin{aligned} &\varphi(0) = 0 \\ &\varphi(1) = 1 \\ &\varphi' > 0 \end{aligned} \tag{15}$$

As a result, what was first perceived as discounting factors, the $\mu_i$'s, now appear as transposed probabilities whose sum is obviously equal to 1.

The combination of a utility function $u$ with decision weights $\mu_i$ (such that $\sum_{i=1}^{n} \mu_i = 1$) determined by a probability weighting function $\varphi$, given by (13) and (14), amounts to what is currently known as "rank-dependent utility"[10]:

$$\begin{aligned} U(L) &= \sum_{i=1}^{n} \mu_i u(x_i) \\ &= \sum_{i=1}^{n} \left(\varphi(G(x_i)) - \varphi(G(x_{i+1}))\right) u(x_i) \\ &= \sum_{i=1}^{n} \left( \left(1 + k(1-q)\frac{1-G(x_i)}{G(x_i)}\right)^{-\frac{1}{1-q}} - \left(1 + k(1-q)\frac{1-G(x_{i+1})}{G(x_{i+1})}\right)^{-\frac{1}{1-q}} \right) u(x_i) \\ &= \sum_{i=1}^{n} \left( \exp_q\left(k\frac{1-G(x_i)}{G(x_i)}\right) - \exp_q\left(k\frac{1-G(x_{i+1})}{G(x_{i+1})}\right) \right) u(x_i) \end{aligned} \tag{16}$$

[9]Note that in the case where $i = n$, the probability of obtaining strictly more than $x_n$ is zero, so that $G(x_{n+1}) = 0$.

[10]Rank-dependent utility continues the pioneering work by James Quiggin (1982). For an introduction focusing on associated risk perceptions see, among others, E. Diecidue and P. Wakker (2001), M. Abdellaoui (2009), and M. Cohen (2015). With some qualifications, more recent versions of prospect theory also belong to this kind of models, at least since Tversky and Kahneman's 1992 paper (see Wakker 2010). In several rank-dependent utility models, $U(L)$ is usually written as the (discrete) Choquet integral $U(L) = \sum_{i=0}^{n-1} \varphi(G(x_{i+1}))(u(x_{i+1}) - u(x_i))$, rather than as its equivalent in (16).

It is well-known that when rank-dependent utility prevails, the acknowledged drawbacks of a direct transformation of each single probability, like this of the probability of success in a 2-issues lottery, (the sum of the decision weights might be different from zero and violation of first degree stochastic dominance might occur) do not hold anymore (see, for instance, M. Abdellaoui 2009). The probability weighting function $\varphi$ possesses the expected properties (see (15)) of decision weights in rank-dependent utility, but its shape is more specific, since it is generated by the whole process of probability discounting. Some consequences of the properties of the probability weighting function are discussed in the following section.

# 3 Attitudes conveyed by probability discounting

The properties of the probability weighting function in (14) are controlled by the two parameters $k$ and $q$. The latter were introduced as a recombination of the parameters $a$ and $b$ used in the transformation of physical into perceived delay (4) and of the discount rate in perceived time $r$ (7), and their main virtue seems to have been of rendering possible an expression of time or probability discounting through $q$-discounting. However, they also support the discussion of the underlying attitudes toward rationality, probability and risk.

## 3.1 *Time-rationality and risk-rationality in probability discounting*

A common way to approach time-rationality and risk-rationality is to agree that they rest, respectively, on the fulfillment of axiomatic properties regarding the underlying preferences: *stationarity* for decision in time, and *independence* for decision in risk[11]. Stationarity and independence enter crucially in the axiomatic basis which make, respectively, preferences in time represented by discounted (exponential) utility, and preferences over random variables (lotteries) represented by expected utility. Both are, in their respective domain, a condition for avoiding preference reversal: stationarity guarantees time-consistency, i.e. the constancy of preferences between two gains at different dates, whether close or remote, provided they are separated by the same interval of time; independence preserves our order of preference between two lotteries, whatever the proportions in which they are combined with a third lottery.

Since the decision weights $\mu_i$ can be viewed equivalently as discounting weights ($\mu_i = \psi(l_i) - \psi(l_{i+1})$; see (10)) or as probability weights ($\mu_i = \varphi(G(x_i)) - \varphi(G(x_{i+1}))$; see (14)), a peculiarity of probability discounting is that the issue of rationality is raised simultaneously in relation to time and in relation to risk.

Now, on the one hand, time-rationality is obtained only when $q$ is tending to 1, which yields exponential discounting (and therefore, stationarity and time-consistency) because the ratio between $\psi$ in (10) and its first derivative is a constant equal to $-k$, so that $\mu_i = \exp(-kl_i) - \exp(-kl_{i+1})$. On the other hand, risk-rationality is a special case of simple hyperbolic discounting like in Herrnstein (1981) or Masur (1987), obtained with $q = 0$ in $\psi$ (10). In this case, $\mu_i = (1 + kl_i)^{-1} - (1 + kl_{i+1})^{-1}$: it occurs with the additional condition that $k = 1$, which makes that $\varphi$ in (14) is such that $\varphi(G(x_i)) = G(x_i)$, whatever $x_i$. As a result, when $q = 0$ and $k = 1$, $\mu_i = p_i$, so that probability discounting has generated expected utility (and hence, independence).

This sheds light on the relationship between time-rationality and risk-rationality generated by the transposition of a decision in risk into a decision in time. When moving from the first to the second, we loose time-rationality if the parameters are such that they preserve risk-rationality. Conversely, if we reach time-rationality, we have to give up risk-rationality. Such a conclusion might seem disturbing, but it should not be overestimated. The simple fact that $\mu_i$ can be understood at the same time as a discount factor and as a probability weight, referring respectively to a specific

[11] Stationarity and independence read as follows. Stationarity: assume $x$ and $y$ are two outcomes respectively available at dates $t_1$ and $t_1 + s$. If $(x, t_1)$ and $(y, t_1 + s)$ are indifferent, $(x, t_2)$ and $(y, t_2 + s)$ are also indifferent for any $t_2 \neq t_1$. Independence: assume three lotteries $L_1$, $L_2$ and $L_3$, and any $\lambda \in [0, 1]$. If $L_1$ is preferred to $L_2$, then $\lambda L_1 + (1 - \lambda) L_3$ is also preferred to $\lambda L_2 + (1 - \lambda) L_3$.

case of generalized hyperbolic discounting (10) and of a probability weighting function in rank-dependent utility (14), means that probability discounting should satisfy the criteria of rationality, obviously weaker, which characterize each of these two approaches: the *Thomsen condition of separability* (Fishburn and Rubinstein 1982, pp. 686-687) for time-rationality[12], and *comonotonic tradeoff consistency* (Wakker 1994, p. 13) for risk-rationality[13]. Taking seriously the idea on which probability discounting is based, namely that deciding in risk might be viewed as a way of deciding in time, entails that something has to be abandoned in our requirements in terms of rationality: either one of the two types of rationality (in time or in risk), when the parameters $k$ and $q$ are given appropriate values or, in the general case, the strong versions of risk-rationality and time-rationality, in favour of the weaker versions consistent with rank-dependent utility and generalized hyperbolic discounting.

## 3.2 The probability discounting determination of attitudes toward probabilities and risk

### 3.2.1 The shape of the probability weighting function

Let us start with the properties of the probability weighting function $\varphi$ defined as in (14). We know that this function is increasing, since its first derivative is positive on $[0,1]$:

$$\varphi'(p) = kp^{-2}\left(1+k(1-q)\frac{1-p}{p}\right)^{-\frac{2-q}{1-q}} > 0 \tag{17}$$

Its second derivative is

$$\varphi''(p) = -kp^{-4}\left(1+k(1-q)\frac{1-p}{p}\right)^{-\frac{2-q}{1-q}}\left(2p-k(2-q)\left(1+k(1-q)\frac{1-p}{p}\right)^{-1}\right) \tag{18}$$

The part played by the parameters $k$ and $q$ is crucial. According to their values, $\varphi$" is either positive, or negative, or of alternate signs, so that $\varphi$ is either convex, or concave, or inverse S-shaped (firstly concave, then convex), or S-shaped (firstly convex then concave).

$\varphi''$ can be rewritten: $\varphi''(p) = A(p) \times B(p)$
where $A(p) = -kp^{-4}\left(1+k(1-q)\frac{1-p}{p}\right)^{-\frac{2-q}{1-q}}$ and $B(p) = \left(2p-k(2-q)\left(1+k(1-q)\frac{1-p}{p}\right)^{-1}\right)$

$A(p)$ is always negative. Hence, the sign of $\varphi''(p)$ depends on the sign of $B(p)$, which writes also:

$$B(p) = \frac{2p(1-k(1-q))-kq}{p+k(1-q)(1-p)} \tag{19}$$

Let us analyse the sign of $B(p)$ with respect to the values of $q$.

**a.** If $q=0$ then replacing $q$ by 0 in $B(p)$ leads to $B(p) = \frac{2p(1-k)}{p+k(1-p)}$. Since the denominator

[12] Thomsen condition of separability (Fishburn and Rubinstein 1982) is based on the idea that when deciding in time, we compensate differences in outcomes by differences in dates, and that these differences are additive. So that given three outcomes $x$, $y$ and $z$ and three dates $r$, $s$ and $t$, if $(x,t)$ and $(y,s)$ are indifferent to a decision-maker, as well as $(y,r)$ and $(z,t)$, it means that $x-y$ is compensated by $t-s$ and, on the other hand, $y-z$ by $r-t$. This means that $x-z=(x-y)+(y-z)$ is compensated by $r-s=(r-t)+(t-s)$. And therefore, $(x,r)$ is also indifferent to $(z,s)$.

[13] Comonotonic tradeoff consistency (Wakker 1994) reads as follows. Assume two sets of pairwise lotteries defined as $L_\alpha = (x_1,...,\alpha,...,x_n;p_1,...,p_i,...,p_n)$, $L_\beta = (y_1,...,\beta,...,y_n;p_1,...,p_i,...,p_n)$ and as $L_\gamma = (x_1,...,\gamma,...,x_n;p_1,...,p_i,...,p_n)$, $L_\delta = (y_1,...,\delta,...,y_n;p_1,...,y_i,...,p_n)$. If for some $i$, there exists outcomes $\alpha,\beta,\gamma,\delta$ so that $L_\alpha$ is preferred to $L_\beta$ and $L_\delta$ is preferred to $L_\gamma$, then for two other alternative sets of lotteries defined in the same way, there is no $i'$ for which $L'_\alpha$ is preferred to $L'_\beta$ and, contrary to the previous case, $L'_\gamma$ is strictly preferred to $L'_\delta$. Alternative key axioms are given by A. Chateauneuf 1999, pp. 25-27.

is always positive, then the sign of $B(p)$ depends on the sign of its numerator. As a consequence, $B(p)$ is positive if and only if $k \leq 1$. Hence when $q = 0$, $\varphi$ is concave if and only if $k \leq 1$, and $\varphi$ is convex otherwise.

**b.** If $q \in ]0, 1[$ then according to equation (19), two cases can occur: the case where $k < \frac{1}{1-q}$ and the case where $k \geq \frac{1}{1-q}$.

- If $k \geq \frac{1}{1-q}$ then $B(p)$ is negative whatever $p \in [0, 1]$. This leads to $\varphi''(p) \geq 0$ ($\varphi(p)$ convex) on the interval $[0, 1]$.

- If $k < \frac{1}{1-q}$ then $B(p)$ is negative (see the numerator of $B(p)$) on the interval $[0, p_0]$ and is positive on the interval $[p_0, +\infty]$, where $p_0 = \frac{kq}{2(1-k(1-q))}$. However $p$ (a probability) cannot go beyond 1. This means that $p_0$ is either less than 1 or higher than 1. $p_0$ is less than 1 if and only if $k < \frac{1}{1-\frac{q}{2}}$. Hence:

  - when $q \in ]0, 1[$, if $k < \frac{1}{1-q}$ and $k < \frac{1}{1-\frac{q}{2}}$ then $B(p)$ is negative on the interval $[0, p_0]$ and positive on the interval $[p_0, 1]$; that is, $\varphi''(p) \geq 0$ ($\varphi(p)$ convex) on the interval $[0, p_0]$ and $\varphi''(p) \leq 0$ ($\varphi(p)$ concave) on the interval $[p_0, 1]$;
  - when $q \in ]0, 1[$, if $k < \frac{1}{1-q}$ and $k \geq \frac{1}{1-\frac{q}{2}}$ then $B(p)$ is negative on the interval $[0, 1]$; that is, $\varphi''(p) \geq 0$ ($\varphi(p)$ convex) on the interval $[0, 1]$.

**c.** If $q < 0$ then according to equation (19), two cases can occur: the case where $k \leq \frac{1}{1-q}$ and the case where $k > \frac{1}{1-q}$.

- If $k \geq \frac{1}{1-q}$ then $B(p)$ is positive whatever $p \in [0, 1]$. This leads to $\varphi''(p) \leq 0$ ($\varphi(p)$ concave) on the interval $[0, 1]$.

- If $k > \frac{1}{1-q}$ then $B(p)$ is positive (see the numerator of $B(p)$) on the interval $[0, p_0]$ and is negative on the interval $[p_0, +\infty]$, where $p_0 = \frac{kq}{2(1-k(1-q))}$. However $p$ (a probability) cannot go beyond 1. This means that $p_0$ is either less than 1 or higher than 1. $p_0$ is less than 1 if and only if $k > \frac{1}{1-\frac{q}{2}}$. Hence:

  - when $q < 0$, if $k > \frac{1}{1-q}$ and $k > \frac{1}{1-\frac{q}{2}}$ then $B(p)$ is positive on the interval $[0, p_0]$ and negative on the interval $[p_0, 1]$; that is, $\varphi''(p) \leq 0$ ($\varphi(p)$ concave) on the interval $[0, p_0]$ and $\varphi''(p) \geq 0$ ($\varphi(p)$ convex) on the interval $[p_0, 1]$;
  - when $q < 0$, if $k > \frac{1}{1-q}$ and $k \leq \frac{1}{1-\frac{q}{2}}$ then $B(p)$ is positive on the interval $[0, 1]$; that is, $\varphi''(p) \leq 0$ ($\varphi(p)$ concave) on the interval $[0, 1]$.

*What are the implications of the above results on the shape and properties of the graph of $\varphi$?*

Since $\varphi(0) = 0$ and $\varphi(1) = 1$, then it is obvious that $\varphi(p) \leq p, \ \forall p$ (respectively : $\varphi(p) \geq p, \ \forall p$) when $\varphi$ is (fully) convex (resp., concave) on the interval $[0, 1]$.

However when $\varphi$ is firstly convex then concave (S-shaped; case with $0 < q < 1$ and $k < \frac{1}{1-\frac{q}{2}}$) or when it is firstly concave then convex (inverse S-shaped; case with $q < 0$ and $k \geq \frac{1}{1-\frac{q}{2}}$), it is not straightforward to conclude whether its graph crosses the first bisector, or whether it does not cross it, because it lies entirely above or under this bisector. The difference between the two situations ($\varphi$ crossing the bisector, or not crossing it) amounts to the existence (in the first situation) or to the non-existence (in the second) of $p^* \in ]0, 1[$ such that

$$\varphi(p^*) = p^* \tag{20}$$

Remind (see (14)) that $\varphi(p) = \left(1 + k\,(1-q)\,\frac{(1-p)}{p}\right)^{-\frac{1}{1-q}}$. Hence equation (20) writes:

$$\left(1 + k\,(1-q)\,\frac{(1-p)}{p}\right)^{-\frac{1}{1-q}} = p \tag{21}$$

Since $\varphi(p)$ is a positive and monotonic function on its domain of definition, (21) writes: $1 + k\,(1-q)\,\frac{(1-p)}{p} = p^{(q-1)}$. That is,

$$\frac{(1-k(1-q))\,p + k\,(1-q) - p^q}{p} = 0$$

As a consequence, we want to know if the below equation (22) admits a root belonging to the interval $]0,1[$ - in which case it crosses the first bisector (otherwise, it does not):

$$-p^q + (1-k(1-q))\,p + k\,(1-q) = 0 \tag{22}$$

Denote $\eta(p) = -p^q + (1-k(1-q))\,p + k\,(1-q)$

- Let us take the case of $\varphi$ S-shaped, with $0 < q < 1$ and $k < \frac{1}{1-\frac{q}{2}}$. We can see that $\eta(0) = k(1-q) > 0$, $\eta(1) = 0$, and $\eta'(p) = -q(p)^{q-1} + (1-k(1-q))$.
  So that $\eta'(p) \geq 0$ if and only if $p \geq \left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}}$.

  - If $k < 1$ then $\left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}} < 1$. As a consequence $\eta$ will decrease on the interval $\left[0, \left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}}\right]$ and will increase on the interval $\left[\left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}}, 1\right]$. However since $\eta(0) > 0$, and $\eta(1) = 0$ then it is necessarily the case that there exist $p^* < \left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}}$ such that $\eta(p^*) = 0$. This proves that when $0 < q < 1$, $k < \frac{1}{1-\frac{q}{2}}$ and $k < 1$, there exists $p^*$ such that $\varphi(p^*) = p^*$. This means that $\varphi$ is S-shaped and crosses the bisector at $p^* < \left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}}$.
  - If $k \geq 1$ then $\left(\frac{q}{1-k(1-q)}\right)^{\frac{1}{1-q}} \geq 1$, and $\eta$ decreases on the interval $[0,1]$. This means that when $0 < q < 1$, $k < \frac{1}{1-\frac{q}{2}}$ and $k \geq 1$, then $\varphi$ is S-shaped and fully under the bisector.

- Likewise if we take the case of $\varphi$ inverse S-shaped, with $q < 0$ and $k \geq \frac{1}{1-\frac{q}{2}}$,

  - if $k > 1$, there exists $p^*$ such that $\varphi(p^*) = p^*$ - i.e. $\varphi$ crosses the bisector;
  - if $k \leq 1$, $\varphi$ is fully above the bisector.

We can therefore write, in the following propositions, the first line after the headers, from which the rest of the tables proceeds (see comments below, in § 3.2.3).

### 3.2.2 Propositions

**Proposition 1** *The table below indicates the link between attitude toward probabilities, attitude toward outcomes and attitude toward risk when the discounting parameter $q$ lies on the interval* $]0,1[$.

**Table 1:** $0 < q < 1$ - Attitudes toward probabilities, outcomes and risk

| $k$ | | 0 | 1 | $\frac{1}{1-\frac{q}{2}}$ $+\infty$ |
|---|---|---|---|---|
| $\varphi$ | | S-shaped, crossing bisector (see Fig. 1) | S-shaped, under bisector (see Fig. 2) | Convex (see Fig. 3) |
| Attitude toward Probability (Strong) | | Local Strong Pessimism and local Strong Optimism (unlikelihood insensitivity) | | Strong Pessimism |
| Attitude toward Probability (Weak) | | Local Weak Pessimism and local Weak Optimism | Weak Pessimism | |
| $u$ concave (decreasing sensitivity) | Attitude toward Risk (Strong) | Neither Strong Risk Averse, nor Strong Risk Seeker | | Strong Risk Averse |
| | Attitude toward Risk (Monotone) | Not Monotone Risk Averse | Monotone Risk Averse | |
| | Attitude toward Risk (Weak) | Not Weak Risk Averse | Weak Risk Averse | |
| $u$ convex (increasing sensitivity) | Attitude toward Risk (Strong) | Neither Strong Risk Averse, nor Strong Risk Seeker | | |
| | Attitude toward Risk (Monotone) | Not Monotone Risk Averse | Monotone Risk Averse when $G_u \leq k$ Not Monotone Risk Averse when $G_u > k$ | |
| | Attitude toward Risk (Weak) | Not Weak Risk Averse | Weak Risk Averse if $G_u \leq k$, or there exists $g \geq 1$ | |

Remarks:
- $G_u = \sup\limits_{y<x} \frac{u'(x)}{u'(y)}$
- $g \geq 1$ is such that $u'(x) \leq g\frac{u(x)-u(y)}{x-y}$, for $x > y$, and $\varphi(p) \leq p^g$

**Proposition 2** *The table below indicates the link between attitude toward probabilities, attitude toward outcomes and attitude toward risk when the discounting parameter $q = 0$.*

**Table 2:** $q = 0$ - Attitudes toward probabilities, outcomes and risk

| $k$ | | 0 | $1 = \frac{1}{1-\frac{q}{2}}$ $+\infty$ |
|---|---|---|---|
| $\varphi$ | | Concave (see Fig. 4) | Convex (see Fig. 5) |
| Attitude toward Probability (Strong) | | Strong Optimism | Strong Pessimism |
| Attitude toward Probability (Weak) | | Weak Optimism | Weak Pessimism |
| $u$ concave (decreasing sensitivity) | Attitude toward Risk (Strong) | Neither Strong Risk Averse, nor Strong Risk Seeker | Strong Risk Averse |
| | Attitude toward Risk (Monotone) | Monotone Risk Seeker if $T_u \leq 1/k$ Not Monotone Risk Seeker if $T_u > 1/k$ | Monotone Risk Averse |
| | Attitude toward Risk (Weak) | Weak Risk Seeker if $T_u \leq 1/k$, or there exists $h \geq 1$ | Weak Risk Averse |
| $u$ convex (increasing sensitivity) | Attitude toward Risk (Strong) | Strong Risk Seeker | Neither Strong Risk Averse, nor Strong Risk Seeker |
| | Attitude toward Risk (Monotone) | Monotone Risk Seeker | Monotone Risk Averse when $G_u \leq k$ Not Monotone Risk Averse if $G_u > k$ |
| | Attitude toward Risk (Weak) | Weak Risk Seeker | Weak Risk Averse if $G_u \leq k$, or there exists $g \geq 1$ |

Remarks:
- $T_u = \sup\limits_{y<x} \frac{u'(y)}{u'(x)}$
- $h \geq 1$ is such that $u'(y) \leq h\frac{u(x)-u(y)}{x-y}$, for $x > y$, and $\varphi(p) \leq 1-(1-p)^h$
- $G_u = \sup\limits_{y<x} \frac{u'(x)}{u'(y)}$
- $g \geq 1$ is such that $u'(x) \leq g\frac{u(x)-u(y)}{x-y}$, for $x > y$, and $\varphi(p) \leq p^g$

**Proposition 3** *The table below indicates the link between attitude toward probabilities, attitude toward outcomes and attitude toward risk when the discounting parameter $q$ is strictly negative.*

**Table 3:** $q < 0$ - Attitudes toward probabilities, outcomes and risk

| $k$ | | $0$ | $\frac{1}{1-\frac{q}{2}}$ | $1$ $+\infty$ |
|---|---|---|---|---|
| $\varphi$ | | Concave (see Fig. 6) | Inverse S-shaped, above bisector (see Fig. 7) | Inverse S-shaped, crossing bisector (see Fig. 8) |
| Attitude toward Probability (Strong) | | Strong Optimism | Local Strong Optimism and local Strong Pessimism (likelihood insensitivity) | |
| Attitude toward Probability (Weak) | | Weak Optimism | | Local Weak Optimism and local Weak Pessimism |
| $u$ concave (decreasing sensitivity) | Attitude toward Risk (Strong) | Neither Strong Risk Averse, nor Strong Risk Seeker | | |
| | Attitude toward Risk (Monotone) | Monotone Risk Seeker if $T_u \leq 1/k$ Not Monotone Risk Seeker if $T_u > 1/k$ | | Not Monotone Risk Seeker |
| | Attitude toward Risk (Weak) | Weak Risk Seeker if $T_u \leq 1/k$, or there exists $h \geq 1$ | | Not Weak Risk Seeker |
| $u$ convex (increasing sensitivity) | Attitude toward Risk (Strong) | Strong Risk Seeker | Neither Strong Risk Averse, nor Strong Risk Seeker | |
| | Attitude toward Risk (Monotone) | Monotone Risk Seeker | | Not Monotone Risk Seeker |
| | Attitude toward Risk (Weak) | Weak Risk Seeker | | Not Weak Risk Seeker |

Remarks:

- $T_u = \sup\limits_{y<x} \frac{u'(y)}{u'(x)}$
- $h \geq 1$ is such that $u'(y) \leq h\frac{u(x)-u(y)}{x-y}$, for $x > y$, and $\varphi(p) \leq 1-(1-p)^h$

### 3.2.3 Comments

**a. On the attitudes toward probabilities**

Drawing on (18), the properties of the probability weighting function in relation to the parameters $q$ and $k$ are listed in the first lines of the tables in Propositions 1, 2 and 3.

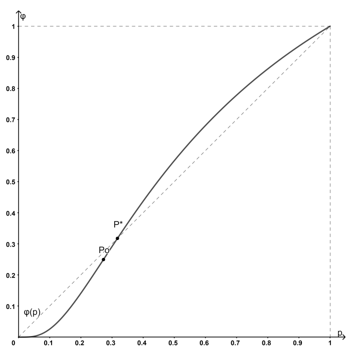

**Figure 1:** Probability weighting function: $\varphi$ S-shaped, crossing the bisector.
$q = .8$, $k = .6$, $p_0 = .27, p* = .32$

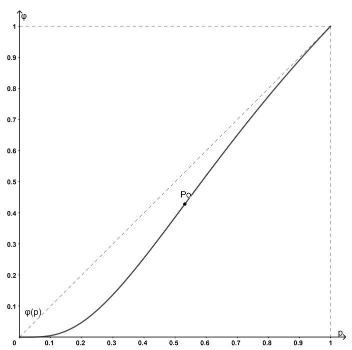

**Figure 2:** Probability weighting function: $\varphi$ S-shaped, under the bisector.
$q = .8$, $k = 1.05$, $p_0 = .53$

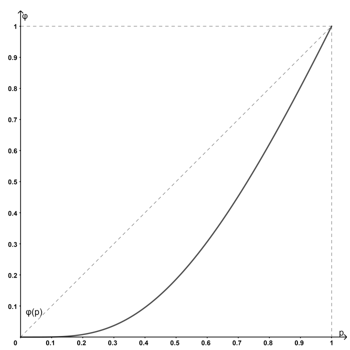

**Figure 3:** Probability weighting function: $\varphi$ convex.
$q = .8$, $k = 2$, $p_0 = 1.33$

The block of lines which follows immediately the header in each table deals with the attitude toward probabilities embodied in the probability weighting function. Generally speaking, it amounts to pessimism or optimism, which can be approached from two different points of view, each one being linked to a way of considering the generic term in the expression of the rank-dependent utility of a lottery: either an utility multiplied by a difference between transformed probabilities, or a transformed probability multiplied by a difference between utilities.

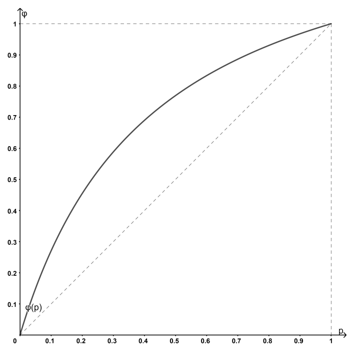

**Figure 4:** Probability weighting function: $\varphi$ concave.
$q = 0$, $k = .3$

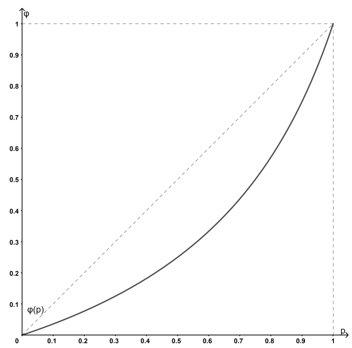

**Figure 5:** Probability weighting function: $\varphi$ convex.
$q = 0, k = 3$

The first point of view (Yaari 1987; Chateauneuf and Cohen 1994) contrasts *strong pessimism* with *strong optimism* (which meets Wakker's 1994 distinction between "probabilistic risk aversion" and "probabilistic risk seeking"), associated respectively to the convexity and to the concavity of $\varphi$. The weight of a typical element $u(x_i)$ in (16) is given by a transformed probability $\mu_i = \varphi(G(x_i)) - \varphi(G(x_{i+1}))$. In particular, a finite variation of $G$ in the neighborhood of 1 or of 0, corresponding to the lowest or to the highest outcomes, indicates its decisional weight $\mu_i$ at the endpoints of the domain of definition of $\varphi$ by the corresponding variation in ordinate. The convexity (fig. 3 or 5) (resp., the concavity (fig. 4 or 6)) of $\varphi$ therefore amounts to strong pessimism (resp., strong optimism), insofar as the probability of the lowest outcomes is overweighted (resp., underweighted), whereas the probability of the highest outcomes is underweighted (resp., overweighted). Strong pessimism (resp., strong optimism) can be interpreted as increasing (resp., decreasing) sensitivity to probability changes when moving from the low probabilities of getting at least the higher outcomes to the high probabilities of getting at least the lower outcomes. This makes easier the interpretation of the intermediate situations of an inverse S-shaped (first concave, then convex; see fig. 7 and 8) or S-shaped (first convex, then concave; see fig. 1 and 2) probability weighting function (see the seminal paper of R. Gonzalez and G. Wu 1999). In the case of an inverse S-shaped function (fig. 7 and 8), the probabilities of the lowest and of the highest outcomes are overweighted relatively to those of the medium outcomes (in the neighbourhood of the inflexion point $p_0$) which are underweighted. This boils down to strong optimism toward medium to high outcomes (the concave part of $\varphi$), and strong pessimism toward low to medium outcomes (its convex part). Commonly used in cumulative prospect theory (see A. Tversky and D. Kahneman 1992), the inverse S-shaped probability weighting function is interpreted in terms of cognitive ability after P. Wakker (see 2010, pp. 203 *sqq*) who called it "likelihood insensitivity", in the sense that people fail to distinguish sufficiently variations of probabilities for medium, usual outcomes, but are overly sensitive when these changes concern best ranked and worst ranked unusual outcomes. Obviously, a symetrical interpretation can be given to the less common S-shaped probability weighting function (fig. 1 and 2), which can be viewed as an expression of what might be called "unlikelihood insensitivity".

The second point of view makes a distinction between what is usually refered to as *weak pessimism* and *weak optimism* (see Cohen 1995). At the difference of strong pessimism and strong optimism, weak pessimism and weak optimism are implicitly based on the interpretation of $\varphi(G(x_i))$ as the transformed probability which we associate to a minimum additional utility $u(x_i) - u(x_{i-1})$ (see *supra*, n. 10, p. 6). In an expected utility framework, we know that $\varphi(G(x_i)) = G(x_i)$ for each $i$. So that pessimism can be seen as doing worse than expected utility, and optimism as doing better than it. Weak pessimism therefore occurs (resp., weak optimism) when $\varphi(G(x_i)) \leq G(x_i)$ (resp., $\varphi(G(x_i)) \geq G(x_i)$), the probabilities of additional utilities being underweighted (resp.,

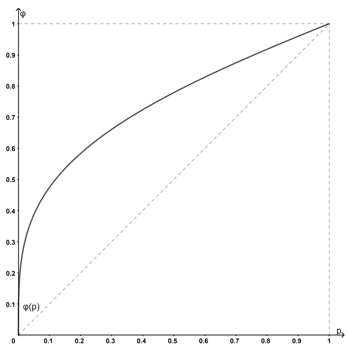


**Figure 6:** Probability weighting function: $\varphi$ concave.
$q = -.5$, $k = .4$, $p_0 = -0.25$

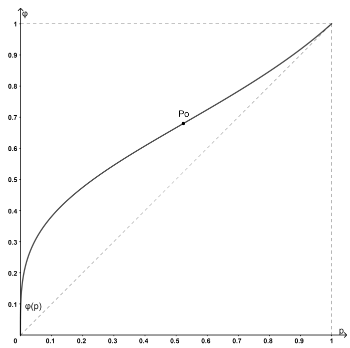


**Figure 7:** Probability weighting function: $\varphi$ inverse S-shaped, above the bisector.
$q = -2.5$, $k = .9$, $p_0 = .52$

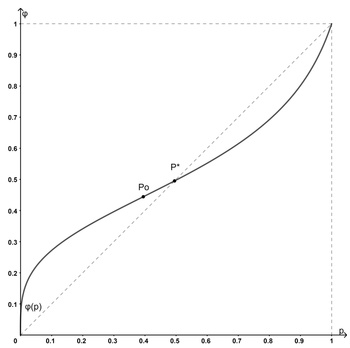


**Figure 8:** Probability weighting function: $\varphi$ inverse S-shaped, crossing the bisector.
$q = -2.5$, $k = 3$, $p_0 = .39, p* = .5$

overweighted). It is obvious that strong pessimism implies weak pessimism (and strong optimism implies weak optimism), whereas the reverse is not true. The previous issue of the convexity or concavity of the probability weighting function is here replaced by the question of knowing whether $\varphi$ lies below the bisector (weak pessimism) (see fig. 2, 3, 5) or above it (weak optimism) (see fig. 4, 6, 7). Consequently, S-shaped or inverse S-shaped probability weighting functions are now significant only when they cross the bisector. When $\varphi$ is inverse S-shaped crossing the bisector (fig. 8), weak optimism prevails locally for relatively high outcomes (with probabilities of winning at least this outcome belonging to the interval between 0 and the abciss $p*$ of the point of intersection of $\varphi$ and the bisector) because the corresponding part of $\varphi$ lies above the bisector; and weak pessimism prevails locally for relatively low outcomes (with probabilities of winning at least this outcome belonging to the interval between $p*$ and 1) because the corresponding part of $\varphi$ lies below the bisector. Of course, an S-shaped $\varphi$ crossing the bisector (fig. 1) is interpreted in a symetrical way.

**b. On the attitudes toward risk**

Following Rothschild and Stiglitz' seminal paper from 1970, we are used to distinguish *weak* and *strong risk-aversion* (resp., *risk-seeking*; *risk-neutrality* being equivalent to risk-aversion and risk-seeking). Both provide answers to different questions. A decision-maker is said to be weakly risk-averse (resp., weakly risk-seeker), if he or she prefers a lottery $L$ to its expected value $E_L(x)$ (resp., the expected value $E_L(x)$ of a lottery $L$ to this lottery). By contrast, a decision-maker is strongly risk-averse (resp., strongly risk-seeker) when, given a pair of lotteries $L_1$ and $L_2$ with equal means such that $L_1$ is stochastically dominating $L_2$ at degree 2[14], $L_1$ (resp., $L_2$) is preferred to $L_2$ (resp., $L_1$). Weak risk attitude is the result of a comparison between a risky distribution and a certain outcome, whereas strong risk attitude denotes a comparison between two risky distributions. An intermediary concept was introduced by Quiggin (1992) in relation to what was to become known as rank-dependent utility: *monotone risk-aversion* (resp., *monotone risk-seeking*) denotes a situation where a decision-maker prefers $L_1$ to a lottery $L_2$ (resp., $L_2$ to $L_1$) when $L_2$ is a monotone increase in risk of $L_1$[15]. Strong, monotone and weak risk attitudes are equivalent in standard expected utility, when the decision weights are equal to the corresponding probabilities, since they all depend on the concavity (risk-aversion) or the convexity (risk-seeking) of the utility function, which incorporates the whole relevant information on the attitude toward risk. Such is the case when $q = 0$ and $k = 1$, so that the decision weights $\mu_i$ are equal to the corresponding probabilities $p_i$. Because his or her behaviour boils down to expected utility when $k = 1$, a simple hyperbolic probability discounter ($q = 0$) who is weakly risk-averse (weakly risk-seeker) is also strongly risk-averse (strongly risk-seeker) and monotonely risk-averse (monotonely risk seeker).

But in all other cases, when the utility of a lottery is given by (16), the properties of the

[14] A lottery $L_1$ (whose cumulative distribution function is $F_1$) is stochastically dominating another lottery $L_2$ (whose cumulative distribution function is $F_2$) at degree 2 when, for all $x$ belonging to $[x_1, x_n]$, $\int_{x_1}^{x} (F_1(s) - F_2(s))\, ds \leq 0$.

[15] $L_2$ is a monotone increase in risk of $L_1$ if $L_2 = L_1 + Z$, with $Z$ being comonotone to $L_1$ and $E(Z) = 0$. On the different concepts of attitude toward risk, see Cohen 1995.

utility function $u$ alone are not sufficient to determine the attitude toward risk: it now depends on the properties of both the utility function $u$ and the probability weighting function $\varphi$. Let us therefore turn to the properties of the utility function. Assume, for sake of simplicity, that it is bi-differentiable, and either concave or convex. The concavity and the convexity of $u$ are currently interpreted as, respectively, a *decreasing sensitivity* and an *increasing sensitivity* to outcomes. In a probability discounting framework like the one of (16), the risk attitude carried on by the utility function can be either reinforced or thwarted by the attitude toward probabilities carried on by the probability weighting function. We rely explicitly on some results concerning rank-dependent utility and adapted to $q$-discounting in order to account for the effects on risk attitude of the interaction between the sensitivity to outcomes ($u$) and the attitude toward probability ($\varphi$).

The *first result* is from Quiggin (1992) and Cohen (1995). It shows that strong risk aversion implies monotone risk aversion which implies weak risk aversion and, in the same way, that strong risk seeking implies monotone risk seeking which implies weak risk seeking. The *second result*, from Chew, Karni and Safra (1987) states on the one hand, that decreasing sensitivity and strong pessimism is equivalent to strong risk aversion, on the other it states that increasing sensitivity and strong optimism is equivalent to strong risk seeking. The *third result* is due to Chateauneuf and Cohen (1994). It highlights the link between weak attitude toward risk and weak attitude toward probability, in the sense that weak risk aversion implies weak pessimism and weak risk seeking implies weak optimism. The *fourth result* is also from Chateauneuf and Cohen (1994). It aims at finding the extent of weak pessimism (resp., weak optimism), which can overcome increasing sensitivity (resp., decreasing sensitivity) so that weak risk aversion (resp., weak risk seeking) is made possible. It states that whatever $x$, $y$, with $x > y$, whatever $p \in [0,1]$, if there exists $g \geq 1$ such that $u'(x) \leq g\frac{u(x)-u(y)}{x-y}$ and $\varphi(p) \leq p^g$, then weak risk aversion is satisfied. Likewise, whatever $x$, $y$, with $x > y$, whatever $p \in [0,1]$, if there exists $h \geq 1$ such that $u'(y) \leq h\frac{u(x)-u(y)}{x-y}$ and $\varphi(p) \geq 1-(1-p)^h$, then weak risk seeking is satisfied. The *fifth result* is from Quiggin (1982, 1992; see also Chateauneuf and Cohen 1994). It says that when $u$ is concave (resp., convex), monotone risk aversion, weak risk aversion and weak pessimism are equivalent (respectively, monotone risk seeking, weak risk seeking and weak optimism are equivalent). Finally the *last result* that we use is due to Chateauneuf, Cohen, Meilijson (2005). It improves Chateauneuf and Cohen (1994) by relying on indexes of pessimism or optimism on the one hand, and on indexes of non-concavity or non convexity of the utility function on the other hand. This result states that monotone risk aversion is equivalent to $G_u \leq P_\varphi$, and monotone risk seeking is equivalent to $T_u \leq O_\varphi$, where $G_u = \sup\limits_{y<x}\frac{u'(x)}{u'(y)}$ is an index of non-concavity ($G_u \geq 1$ and is equal to 1 when $u$ is concave), $T_u = \sup\limits_{y<x}\frac{u'(y)}{u'(x)}$ is an index of non-convexity ($T_u \geq 1$ and is equal to 1 when $u$ is convex), $P_\varphi = \inf\limits_{0<p<1}\frac{\frac{1-\varphi(p)}{\varphi(p)}}{\frac{1-p}{p}} \geq 1$ is an index of pessimism, and $O_\varphi = \inf\limits_{0<p<1}\frac{\frac{\varphi(p)}{1-\varphi(p)}}{\frac{p}{1-p}}$ an index of optimism. The result of Chateauneuf, Cohen and Meilijson (2005) therefore expresses situations where pessimism (resp. optimism) compensates the convexity (resp. concavity) of the utility function. It can be shown that when $q$-discounting occurs, $P_\varphi = k$ and $O_\varphi = 1/k$, which are both obtained when $p$ tends to 1. So that the result of Chateauneuf, Cohen and Meilijson (2005) can be reformulated as:

$$\begin{cases}\text{Monotone Risk Aversion} \Leftrightarrow G_u \leq k \\ \text{Monotone Risk Seeking} \Leftrightarrow T_u \leq 1/k\end{cases}$$

Drawing on the above results from the literature, it it has become possible to determine, in Propositions 1, 2 and 3, the various types of attitudes toward risk generated by the combination between an attitude toward probabilities expressed by the properties of $\varphi$, and an attitude toward output which comes from the properties of $u$.

It is commonsense to claim that all this depends on the action of the two parameters, $k$ and $q$. In the case of decision in time, their respective function seems rather clear (see, for instance, Takahashi 2007b, pp. 639-640 and Munoz Torrecillas *et al.* 2018, pp. 191-192). $k$ is usually perceived as a parameter of "impulsivity", which we can understand as "impatience", since it increases the discounting weight of physical waiting time. And $q$ is a parameter of (time-) consistency, since when it moves away from 1, it also makes exponential discounting more and more distant. Regarding decision in risk, $q$ separates situations of non-optimism (in which global risk-seeking of any type is impossible) when it is greater than 0 and smaller than 1 (Proposition 1) from situations of non-pessimism (in which global risk-aversion, also of any type, is impossible) when it is less than 0 (Proposition 3). Rather than a parameter of "risk-aversion", as Takahashi *et al.* (2013, p. 877) first called it, $k$ plays a crucial part as a sophisticated parameter of pessimism: it constitutes the upper-bound for the index of non-concavity $G_u$ in order to obtain monotone risk-aversion; or it represents, through $1/k$, the upper-bound for the index of non-convexity $T_u$ to produce monotone risk-seeking. This shows that appropriate values of $k$ can compensate either the concavity or the convexity of the utility function to produce either monotone risk-seeking in the first case, or monotone risk-aversion in the second case. And if $k$ is either too large or too small for this, it remains possible to have at least sufficient conditions to obtain weak risk-aversion or weak-risk seeking (Chateauneuf and Cohen 1994). When it is smaller than 1 (when $0 < q < 1$) or greater than 1 (when $q < 0$), $k$ generates S-shaped or inverse S-shaped probability weighting functions $\varphi$ which cross the bisector, so that none of the basic *global* attitudes toward risk can exist. In all other cases, at least weak optimism or weak pessimism occurs, so that the necessary condition for any conception of risk aversion or risk seeking is satisfied (Chateauneuf and Cohen 1994). At last, the relation between both parameters, $k$ and $q$, allows determining the range of their relative values for which strong risk attitudes are possible: if $q$ lies between 0 and 1, $k \geq \frac{1}{1-\frac{q}{2}}$ generates strong pessimism, thus determining strong risk-aversion with $u$ concave; symmetrically, if $q$ is less than 0, $k \leq \frac{1}{1-\frac{q}{2}}$ generates strong optimimism, and strong risk-seeking with $u$ convex (Chew, Karni and Safra 1987).

# 4 Concluding remarks

Emerging from the intuition that probability entails a more or less long delay before winning, probability discounting has shown fruitful. Though usually avoiding the use of an explicit utility function, it could integrate it and give rise to a more complete representation of risky choices. Originally presented in the framework of 2-issues lotteries, its cautiousless extension to the case of $n$-issues lotteries would face the today well-known drawbacks associated to a one-to-one transformation of probabilities, like the violation of stochastic dominance of degree 1. This is why the same kind of transformation as the one in use for rank-dependent utility has been employed. The transformation therefore concerns not a single delay or a single probability before winning, but the average delay before obtaining at least a certain reward, or the (decumulated) probability of getting at least this reward. The effects of this transformation on the rationality of behaviour and on the attitude towards risk depend on the shape of the $q$-discounting function, which applies to both time and probability.

An immediate conclusion can be drawn regarding rationality both in time and in risk. Whereas appropriate values of the parameters of the $q$-discounting function allow reaching the standard criteria of time-rationality (stationarity, through exponential discounted utility) and risk-rationality (independence, through expected utility), they cannot be fulfilled together, the latter being a particular case of hyperbolic discounted utility. The attitude toward risk depends on both the attitude toward outcomes, embedded in the utility function, and on the attitude toward probabilities expressed in extended probability discounting. In a trivial way, the concavity or convexity of the utility function brings respectively risk-aversion or risk-seeking. But these have to be combined with the attitude toward probabilities shown by the $q$-discounting function in a rank-dependent utility framework. Now, according to the values of its parameters, we obtain the whole range of the properties of the probability weighting function usually acknowledged in the literature. This allows distinguishing between the different types (weak and strong) of pessimism and optimism toward probability, and to rely on the few results in the existing literature in order to determine the various attitudes toward risk generated by the combination of a utility function and probability discounting.

Over the last thirty years or so, probability discounting has shown that in a large variety of cases it is an experimentally relevant procedure to account for behaviour under risk. From a theoretical point of view, its generalisation leads to extending its scope and clarifying its meaning in terms of rationality and attitude toward risk.

# References


M. Abdellaoui (2009). Rank-dependent Utility, in P. Anand, P.K. Pattanaik and C. Puppe (eds), *Rational and Social Choice: an Overview of New Foundations and Applications*, Oxford and New York: Oxford University Press, pp. 69-89.

A. Chateauneuf (1999). Comonotonicity Axioms and Rank-Dependent Expected Utility Theory for Arbitrary Consequences, *Journal of Mathematical Economics*, 32, pp. 21-45.

A. Chateauneuf and M. Cohen (1994). Risk Seeking with Diminishing Marginal Utility in a Non-Expected Utility Model, *Journal of Risk and Uncertainty*, 9, pp. 77-91.

A. Chateauneuf, M. Cohen and I. Meilijson (2005). More Pessimism than Greediness: a Characterization of Monotone Risk Aversion in the Rank-Dependent Expected Utility Model, *Economic Theory*, 25, pp. 649-667.

S.-H. Chew, E. Karni and Z. Safra (1987). Risk Aversion in the Theory of Expected Utility with Rank Dependent Preferences, *Journal of Economic Theory*, 42, pp. 370-381.

M. Cohen (1995). Risk-Aversion Concepts in Expected- and Non-Expected-Utility Models, *Geneva Papers on Risk and Insurance Theory*, 20, pp. 73-91.

M. Cohen (2015). Risk Perception, Risk Attitude and Decision: a Rank-Dependent Approach, *Mathematical Population Studies*, 22(1), pp. 53-70.

S. Dehaene (2003). The Neural Basis of the Weber–Fechner Law: a Logarithmic Mental Number Line, *TRENDS in Cognitive Sciences*, 7(4), pp. 145-147.

E. Diecidue and P.P. Wakker (2001). On the Intuition of Rank-Dependent Utility, *Journal of Risk and Uncertainty*, 23(3), pp. 281–298.

R. Gonzales and G. Wu (1999). On the Shape of the Probability Weighting Function, *Cognitive Psychology*, 38, pp. 129-166.

L. Green and J. Myerson (2004). A Discounting Framework for Choice With Delayed and Probabilistic Rewards, *Psychological Bulletin*, 130(5), pp. 769-792.

R. Herrnstein (1981). Self-Control as Response Strength. In C.M. Bradshaw, E. Szabadi and C.F. Lowe (eds), *Quantification of Steady-State Operant Behavior*, Amsterdam: Elsevier/North-Holland.

D. Kahneman and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2), pp. 263-291.

G. Loewenstein and D. Prelec (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation, *Quarterly journal of economics*, 107(2), pp. 573-597.

J.E. Mazur (1987). An Adjusting Procedure for Studying Delayed Reinforcement. In M.L. Commons, J.E. Mazur, J. A. Nevin, and H. Rachlin (eds.), *Quantitative Analyses of Behavior (Vol. 5), The Effect of Delay and of Intervening Events on Reinforcement Value*, Hillsdale, NJ: Erlbaum, pp. 55-73.

M.J. Munoz Torrecillas, T. Takahashi, J. Gil Roales-Nieto, S. Cruz Rambaud, Z. Callejon Ruiz and B. Torrecillas Jover (2018). Impatience and Inconsistency in Intertemporal Choice: An Experimental Analysis, *Journal of Behavioral Finance*, 19(2), pp. 190-198.

P. Ostaszewski, L. Green, J. Myerson (1998). Effects of Inflation on the Subjective Value of Delayed and Probabilistic Rewards, *Psychonomic Bulletin and Review*, 5(2), pp. 324-333.

D. Prelec and G. Loewenstein (1991). Decision Making Over Time and Under Uncertainty: A Common Approach, *Management Science*, 37(7), pp. 770-786.

J.C. Quiggin (1982). A Theory of Anticipated Utility, *Journal of Economic Behavior and Organization*, 3, pp. 324-343.

J.C. Quiggin (1992). Increasing Risk: Another Definition, in A. Chikan (ed.), *Progress in Decision, Utility and Risk Theory*, Dordrecht: Kluwer, pp. 239–248.

H. Rachlin, J. Brown and D. Cross (2000). Discounting in Judgments of Delay and Probability, *Journal of Behavioral Decision Making*, 13, pp. 145-159.

H. Rachlin, A.W. Logue, J. Gibbon and M. Frankel (1986). Cognition and Behavior in Studies of Choice, *Psychological Review*, 93(1), pp. 33-45.

H. Rachlin, A. Raineri and D. Cross (1991). Subjective Probability and Delay, *Journal of Experimental Analysis of Behavior*, 55(2), pp. 233-244.

H. Rachlin and E. Siegel (1994). Temporal Patterning in Probabilistic Choice, *Organisational Behavior and Decision Processes*, 59, pp. 161-176.

H. Rachlin, E. Siegel and D. Cross (1994). Lotteries and the Time Horizon, *Psychological Science*, 5(6), pp. 390-393.

M. Rothschild and J. E. Stiglitz (1970). Increasing Risk: I. A Definition, *Journal of Economic Theory*, 2, pp. 225-243.

J.B. Rotter (1954). *Social Learning and Clinical Psychology*, Englewood Cliffs, NJ: Prentice-Hall.

S.S. Stevens (1957). On the Psychological Law, *Psychological Review*, 64(3), pp. 153-181.

T. Takahashi (2005). Loss of Self-Control in Intertemporal Choice May Be Attributable to Logarithmic Time-Perception, *Medical Hypotheses*, 65, pp. 691-693.

T. Takahashi (2007a). A Probabilistic Choice Model Based on Tsallis' Statistics, *Physica A: Statistical Mechanics and its Applications*, 386(1), pp. 335-338.

T. Takahashi (2007b). A Comparison of Intertemporal Choices for Oneself versus Someone Else Based on Tsallis' Statistics, *Physica A: Statistical Mechanics and Its Applications*, 385(2), pp. 637-644.

T. Takahashi (2008). A Comparison Between Tsallis's Statistics-Based and Generalized Quasi-Hyperbolic Discount Models in Humans, *Physica A: Statistical Mechanics and Its Applications*, 387(2-3), pp. 551-556.

T. Takahashi (2010). A Social Discounting Model Based on Tsallis' Statistics, *Physica A: Statistical Mechanics and Its Applications*, 389, pp. 3600-3603.

T. Takahashi (2011). Psychophysics of the Probability Weighting Function, *Physica A: Statistical Mechanics and Its Applications*, 390(5), pp. 902-905.

T. Takahashi (2013). The q-Exponential Social Discounting Functions of Gain and Loss, *Applied Mathematics*, 4, pp. 445-448.

T. Takahashi, R. Han, F. Nakamura (2012). Time Discounting: Psychophysics of Intertemporal and Probabilistic Choices, *Journal of Behavioral Economics and Finance*, 5, pp. 10-14.

T. Takahashi, R. Han, H. Nishinaka, T. Makino, H. Fukui (2013). The $q$-Exponential Probability Discounting of Gain and Loss, *Applied Mathematics*, 4, pp. 876-881.

T. Takahashi, H. Oono and M. H. B. Radford (2007). Comparison of Probabilistic Choice Models in Humans, *Behavioral and Brain Functions*, 3(20) (open access at http://www.behavioralandbrainfunctions.com/content/3/1/20).

C. Tsallis (1994). What Are the Numbers that Experiments Provide?, *Quimica Nova*, 17(6), pp. 468-471.

A. Tversky and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representations of Uncertainty, *Journal of Risk and Uncertainty*, 5, pp. 297-323.

A. Tversky and P. Wakker (1995). Risk Attitudes and Decision Weights, *Econometrica*, 63(6), pp. 1255-1280.

P.P. Wakker (1994). Separating Marginal Utility and Probabilistic Risk Aversion, *Theory and Decision*, 36, pp. 1-44.

P.P. Wakker (2010). *Prospect Theory - For Risk and Ambiguity*, Cambridge: Cambridge University Press.

M.E. Yaari (1987). The Dual Theory of Choice under Risk, *Econometrica*, 55, pp. 95-105.

R. Yi, X. de la Piedad, W.K. Bickel (2006). The Combined Effects of Delay and Probability in Discounting, *Behavioural Processes*, 73, pp. 149-155.