HYPERBOLIC DISCOUNTING AND ECONOMIC POLICY

Benedek Nagy

University of Szeged, Hungary, 6722 Szeged, Kálvária Av. 1., Nagy.Benedek@eco.u-szeged.hu

Abstract: Economic policy-making often entails comparison between immediate costs and flows of future benefits or immediate benefits and series of future costs. Economics has a tool to handle such comparisons: the present- and future value calculations and the net present value rule. Experimental economics, however, has strongly criticised the method of exponential discounting applied in such calculations. Based on experiments for the sake of more psychological realism, they propose alternative methods to the exponential model: hyperbolic and quasi-hyperbolic discounting models.

The present paper has a twofold objective: first, to review these different models and the relationships between them to show how the different models will yield different results when calculating and comparing present values of a single future payment, but even more if we compare present values of flows of future payments. The literature has not yet employed the hyperbolic and quasi-hyperbolic models for such calculations. Second, I point out why it is important to heed the findings of experimental economics especially in the field of economic policy-making.

Key words: hyperbolic discounting, present value calculation, annuity

JEL Classification: D03, D78, D92

DOI: 10.2478/v10135-009-0010-7

HYPERBOLIC DISCOUNTING AND ECONOMIC POLICY[1]

Benedek Nagy[2]

Introduction

It is a commonly used and simple decision rule in discounted cash flow analysis that given two mutually exclusive investment projects, the one yielding a higher positive net present value is to be carried out (net present value being the sum of all the net cash flows in connection with the given project discounted to present). Investment projects can, however, generate external benefits too over multiple periods, the discounted values of which can also be incorporated to give a measure of an overall welfare effect of the project. The primary issue I am going to investigate in this paper concerns the *method* of discounting: if we acknowledge the findings of experimental economics, then instead of conventional exponential discounting in certain situations it might be more appropriate to apply a more psychologically adequate discounting model. I will argue that this can be the case with the external benefits of investment projects envisaged above.

In the first section of this paper I review the discounting models devised to account for time preference, discuss their connections, and calculate the difference between the results obtained with using other models for present value calculations. Based on the results, in the second section I show that the differences arising from using different models will be even more emphatic when we apply these models for calculating present values of *annuities* or flows of yields. To my knowledge the alternative models competing with the exponential model have not so far been applied to this field. As in the first section, here too I present the differences arising from the application of the other models. In the third section I show that while it may be appropriate to continue using the exponential model in certain situations, in others the alternative models may prove more appropriate. From a firm's point of view, when profit maximising in the strict sense is necessary, it is reasonable to use the exponential model to calculate the present value of a given investment's revenues and costs of different time periods, since by depositing the resulting net cash flows in a bank, the firm will receive exactly this much compensation, or by using its own funds will forego exactly this amount of interest. This is the basis for the NPV rule. From a consumer's point of view, however, other motives than the plain monetary ones might also play an important role. If a given decision yields not only material (i.e. monetary), but also other kinds of welfare yields (like utility or consumer surplus), it is reasonable to study how the concerned parties discount these mentally realising or forgone yields, which can not technically be transferred between time periods. In such cases the government might play an important role as a regulator. In section four I explore possible fields of application for this idea to

[2] University of Szeged, Hungary, 6722 Szeged, Kálvária Av. 1., Nagy.Benedek@eco.u-szeged.hu

find the government is involved in decisions that have to take into account the differences arising from the use of non-exponential discounting. Section five offers conclusions.

1. Alternatives to the exponential discounting model

I am going to examine to what degree the use of different discounting methods influences the present value calculations of single utility yields.

The discounted utility model appeared in economic thinking with Paul Samuelson's 1937 paper "A Note on Measurement of Utility", and it quickly gained popularity as the method for comparing utilities arising at different points in time. Samuelson writes: "[D]uring any specified period of time, the individual behaves so as to maximise the sum of all future utilities, they being reduced to comparable magnitudes by suitable time discounting. [...] The individual discounts future utilities in some simple regular fashion which is known to us" (Samuelson 1937, p. 156). The "regular fashion known to us" he mentions is *exponential discounting*.

Rachlin (2006, p. 425) quotes the entry of the *Oxford Encyclopaedic Dictionary* explaining the broad meaning of discounting in the original vein of Samuelson: to "reduce the effect (of an event, etc.) by previous action". In general, consequently, we can speak of an effect of an initial action (X) subsequently lessened (x) by some coefficient (δ). This lessening can be expressed either in the formula $x = \delta X$, or in the formula $\dfrac{x}{X} = \delta$. The coefficient δ itself can be smaller or greater depending on other variables, expressing in response to what and to what degree the effect of the initial action is lessened. In the exponential discounting model this coefficient depends on the interest rate, as in the formula $\delta = \dfrac{1}{1+r}$.

Although Samuelson introduced discounting originally to compare future *utilities*, microeconomics subsequently used it in discounted cash flow analysis, as its name also shows, in a more narrow sense to compare future *cash flows*. In this case X is a sum of money receivable in the future (or the utility thereof), x is a sum of money receivable in the present (or the utility thereof), and δ is a coefficient, the discount factor, using which the value of the above two is equivalent for the decision maker. Now the exact value of δ is determined by two factors: The rate of interest as an exogenous parameter, and the elapsed time as a variable (that is, $\delta = \delta_t$). The more time elapses between the present and the future, the smaller δ_t will be, so a given present sum of money will correspond to a larger sum of money in the future, or a given future sum of money to a smaller present one. In the case of the exponential model we can calculate the present value of a given future sum of money using the two determining factor in the formula $\delta_t = \dfrac{1}{(1+r)^t}$.

Conventional intertemporal choice theory treats the discount factor exponentially, which means that the ratio of x to X changes at a constant rate per unit of time. This exponential discounting will result in preferences consistent in time: what is valued

higher at a given point in time will at any later point of time be valued higher. Experimental economics shows that human subjects in experiments of actual decisions repeatedly make decisions irreconcilable with this: their preferences are dynamically inconsistent. The most famous experimentally studied anomaly resulting from dynamically inconsistent preferences is the notion of preference reversal (Kirby–Herrnstein 1995). The problem of procrastination arises in connection with this (Laibson 1997): This year I think that based on rational calculations it is reasonable for me to start saving heavily next year, but when next year actually comes around, I decide not to, and postpone it one more year, not foreseeing that I will act the same way one year from now. Many fields of economics explore the consequences of dynamically inconsistent preferences from under-saving and inefficient saving for retirement to over-consumption and the exhaustion of non-renewable resources. These kinds of behaviour cannot be explained with exponential discounting, but they do make sense when assuming a different discounting behaviour.

Multiple alternative explanations were devised to explain such phenomena. One group of these, the one I am going to investigate here, *changes the underlying discounting model*. Such is hyperbolic discounting or quasi-hyperbolic – also known as beta-delta – discounting (Phelps–Pollak 1968). These alternative models account for dynamic inconsistency by not only changing the discount factor depending on the time elapsed (decreasing it constantly), but by allowing time to influence the rate of this change as well (basically, the discount rate). These alternative models try not to part drastically with the model previously applied, but rather modify and generalise it.[1]

Several generalisations exist that try to explain discount rate decreasing in time. One of these generalises the discount factor of a given point in time of the exponential model from $\delta_t = \delta^t$ to a more general $\delta_t = \beta\delta^t$ – this is the basic idea of *discrete-time quasi-hyperbolic discounting*. According to this view the difference is attributable to the especially heavy discounting for the first period. Another generalisation of the exponential model yields the formula $\delta_t = \delta^{\alpha(t)}$ – this is the group of *continuous-time hyperbolic discount functions*. This view explains the differences by "distorted time apprehension".[2]

Many experiments justify the usage of such alternative models. Thaler (1981) uses experiments to show that discount rates are changing over time. He asked subjects to state what amount of money receivable in one month/one year/ten years they consider equivalent to a given amount of money now. Based on the mean answers he calculated

---

[1] Many alternative models exist, of which I only use a group giving explicit functional forms to their discounting model. The alternative explanations are thoroughly reviewed in Frederick et al. (2002). This paper also mentions Subadditive discounting which is partly mathematised, but does not give an explicit functional form for the discounting behaviour, in contrast to the models I am using (this is why I do not study this one in more details: for more on subadditive discounting, see Read 2001). Trope and Liberman (2003) mentions alternative explanations for the notion of preference reversal that use psychological factors such as attitudes, emotions and cognition, but these are even less mathematised.

[2] Right or distorted apprehension of time are concepts that inspired many philosophers. I certainly do not wish to tackle the philosophical interpretations of this question, but I will later return to how the economist uses this term.
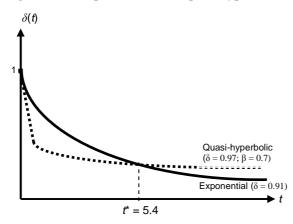
required discount rates per annum, and got 345% for the one month condition and only 19% for the ten years condition. Instead of experiments, Laibson et al. (2007) use field data from actual choices to estimate the discount function and conclude that a quasi-hyperbolic function fits the data better than an exponential one with a short run discount rate of 39.5% and a long run discount rate of 4.3%. Loewenstein–Prelec (1992) derives the generalised hyperbolic discount function theoretically deduced from the very fact of preference reversal. Below I only present the possible consequences in economic applications that can arise if we have reason to assume the decision makers employ one of the above alternative discounting models. I do not wish to take sides in the question whether economic decision makers in fact use one or the other model or which describes their discounting behaviour better but rather what difference it makes if they use the one or the other.

1.1 Discrete-time Quasi-hyperbolic discounting

As mentioned above, quasi-hyperbolic discounting was first used by Phelps and Pollak in their 1968 paper. Laibson (1996) takes over the functional form used by them, which says that the discount factor for the $t = 0$ time period is one, and then the discount factor for any time $t > 0$ can be calculated as $\delta_t = \beta \cdot \delta^t$. If $\beta = 1$, we have exponential discounting as a special case, but if $1 > \beta > 0$, then the resulting discrete-time discount function with values {1, $\beta\delta$, $\beta\delta^2$, $\beta\delta^3$,…} "mimics the qualitative property of the hyperbolic discount function, while maintaining most of the analytical tractability of the exponential discount function" (p. 8). Such considerations must have led other authors when they used laboratory experiments or data from actual decisions to estimate the parameters of the beta-delta discount function. The underlying assumption behind this kind of discount function is that decision makers first discriminate between instant versus delayed consumption, and only then do they look at how long the consumption is delayed. Any non-instantaneous consumption is immediately discounted heavily by a factor of β, but once this one-time heavy discounting is done for consuming later, the actual length of the delay does not matter very much. Delta shows the extent of long run patience (or impatience), and beta and delta together show the extent of short run patience (impatience). In the exponential case beta would be one, and so the decision maker would be equally (to the extent shown by delta) patient or impatient on the short and the long run. A beta smaller than one implies stronger impatience on the short run: the decision maker is more impatient (less patient) on the short run the smaller the value of beta (relative to exponential discounting). Long run (im)patience, however, is shown by delta: the nearer the value of delta is to one in quasi-hyperbolic discounting, the more patient (or less impatient) the decision maker is, again, relative to the exponential model. Contrasting the exponential discount function with the quasi-hyperbolic, we get Figure 1 below. For the sake of easier handling I connected the points of the otherwise discrete quasi-hyperbolic function with the dotted line. The figure shows that given the parameters there exists a t^* point in time when exponential and quasi-hyperbolic discount factors are equal.

Figure 1: An exponential and a quasi-hyperbolic discount function

Source: own processing

According to the figure, if the exponential model's rate of interest is 10%, viz. $\delta_{exp} = 0.91$, while in the quasi-hyperbolic case $\beta = 0.7$ and, due to a smaller long run interest rate, $\delta_{hyp} = 0.97$, then this t^* point in time is at 5.4 years.[1] This means that if someone who mentally discounts the future quasi-hyperbolically with the above parameters put a sum of money in the bank with a shorter than 5.4 years maturity, he/she would now value the amount of money receivable at the end less than the money he/she had to deposit in the bank now. With a maturity longer than 5.4 years, however, the exponentially compounding money he/she would receive at the end would seem more appealing to the decision maker than the money he/she had to deposit now.[2] This decision maker in the first 5.4 years would therefore require more future money for a given present money, or considers less present money to be equivalent to a given present sum, than the bank. *For a period shorter than this the bank would undercompensate the saver, for a period longer than this the bank would overcompensate.*

The above results can be obtained equating the discount factors of the different discounting methods, getting

$$t^* = \frac{\ln \beta}{\ln \delta_{\text{exp}} - \ln \delta_{hyp}}.$$

---

[1] When calculating from the formula at what point of time the different discount factors are equal we can arrive at fractions like in this case, whereas earlier I introduced this kind of generalisation as a discrete-time function. Keeping this in mind I continue to give and use fractional solutions.

[2] It is important to emphasise that these comparisons hold only in the present, because of the above mentioned important (but in this case disturbing) notion of dynamic inconsistency associated with non-exponential discounting. The same comparison made one day later may yield the opposite result. Hereafter I am going to use this meaning when I say decision-makers value incomes or flows of incomes differently when discounted exponentially or non-exponentially.

From the formula we can see that $\dfrac{\partial t*}{\partial \beta} > 0$, and $\dfrac{\partial t*}{\partial r_{\exp}} < 0$; moreover, $\dfrac{\partial t*}{\partial r_{hyp}} > 0$,[1]

and since $r_{\exp} > r_{hyp}$, $t*$ also increases with the decrease of the difference between the two discount rates. In section 2 we will see what further relevance this $t*$ value has. Table 1 shows the t* values for various combinations of parameters. In the rows we have δ and in the columns β parameters for the quasi-hyperbolic discounting following Laibson (1996).

Table 1a: $t*$ values at $\delta_{exp} = 0.952$ (i.e.: $r_{exp} = 0.05$)

$\delta\ hyp \setminus \beta$	0,25	0,5	0,75	0,8
0,96	173,98	86,99	36,10	28,00
0,97	75,63	37,81	15,69	12,17
0,99	35,78	17,86	7,43	5,76

Source: own computation

Table 1b: $t*$ values at $\delta_{exp} = 0.909$ (i.e.: $r_{exp} = 0.1$)

$\delta\ hyp \setminus \beta$	0,25	0,5	0,75	0,8
0,96	25,44	12,72	5,28	4,10
0,97	21,38	10,69	4,44	3,44
0,99	16,26	8,13	3,37	2,62

Source: own computation

The main drawback of the beta-delta discounting model is that it can not be interpreted in continuous-time. Even though I sloppily drew the quasi-hyperbolic discount function as continuous on Figure 1, at least between the 0th and the 1st time periods it is not even theoretically clear how to make it continuous. This major problem is solved by a different kind of generalisation of the exponential model.

1.2. Continuous-time hyperbolic discounting

Cairns (2006) quotes the comparison of Albrecht and Weber between different hyperbolic models. The conventional model can be generalised by the following formula:

$$\frac{x}{X} = \delta_t = \frac{1}{(1+r)^{\alpha(t)}}$$

This generalisation states that the discount factor depends on the interest rate and time; not, however, the real progress of time, but on how the progress of time is apprehended – this is the factor in the formula denoted by $\alpha(t)$.[2]

---

[1] Recall that $\delta_{\exp} = \dfrac{1}{1 + r_{\exp}}$ and $\delta_{hyp} = \dfrac{1}{1 + r_{hyp}}$.

[2] Albrecht and Weber call this δ_t factor *decision weight*, implying that discount factor plays the same role in intertemporal decisions as occurrence probability in the theory of risky decisions.

In the exponential model economic actors apprehend the progress of time as it in fact progresses. In this case $\alpha(t) = t$. We have no more reason to assume that economic actors behave this way than to assume that they do not, but this is a good benchmark assumption; all the more so because banks use this formula when calculating interest and compound interest, and give people exactly that much compensation for their patience as they would require if they apprehended the progress of time right. If an economic actor has a different apprehension of the progress of time (mentally), then the interest promised by a bank will over- or undercompensate him/her for his/her patience. Hyperbolic discounting models are examples of such misapprehension of time, inasmuch as they are special cases of the above formula.[1]

In this paper I examine three kinds of alternative hyperbolic discounting models. The first is the model of Loewenstein and Prelec (1992), the second that of Mazur (1987) and the third that of Harvey (1995).

The discount factor in Loewenstein and Prelec's model is

$$\delta_t = \frac{1}{(1+gt)^{h/g}} \text{ , which we get by setting } \alpha(t) = \frac{h\ln(1+gt)}{g\ln(1+r)},$$

the Mazur model's discount factor is

$$\delta_t = \frac{1}{1+gt} \text{ , with the appropriate substitution } \alpha(t) = \frac{\ln(1+gt)}{\ln(1+r)},$$

and according to Harvey, the discount factor is

$$\delta_t = \frac{1}{(1+t)^h} \text{ , when } \alpha(t) = h\frac{\ln(1+t)}{\ln(1+r)}.$$

It can easily be shown that all the above discount factors increase in t, so the more time elapses in reality, the decision makers will also feel more time has elapsed. All of the above discount factors will give 1 at $t = 0$; that is, whichever model we use, the present is given the same weight. If we introduce the notion of instantaneous discount rate at time t, calculated as $-\dfrac{\delta_t'}{\delta_t}$, then we can grasp growing impatience: this rate is independent of t in the exponential case, while it decreases in t in all the other hyperbolic cases (Laibson 1996).[2]
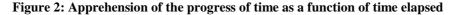
---

[1] Although $\alpha(t)$ was introduced as a distortion in time-apprehension below the formula for $\alpha(t)$ in the Loewenstein–Prelec, the Mazur and Harvey models will show that it is determined among other things by an economic factor external to the decision-maker: the rate of interest. From a psychological point of view it is quite unacceptable that (for example) a grandmother would apprehend the progress of time slower or faster *only because* the interest rate changes. It might be more appropriate thus to refer to $\alpha(t)$ as a factor describing how we should distort the time apprehension of an exponential discounter to act like the hyperbolic discounter in question. Henceforth I am going to use the term "distorted time apprehension" in this sense.
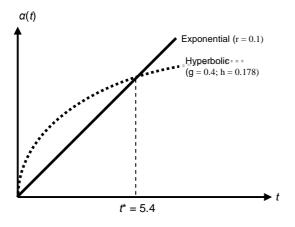
[2] Instantaneous discount rate at time t will be $\ln(1/\delta)$ in the hyperbolic case, and $h / (1 + gt)$ in the

Among the above formulae, the Loewenstein–Prelec model is the most general, all the others can be obtained from that by setting the parameters appropriately. For this reason my statements below will be made only for this model, but with appropriate substitutions they are valid for the other models as well.

Hyperbolic models use the following apprehension of time, contrasted with the exponential case (Figure 2.):

Figure 2: Apprehension of the progress of time as a function of time elapsed

Source: own processing

Again we can determine the t^* point in time by solving $\alpha(t) = t$. Since these functions are defined continuously, interpreting a fractional result is not problematic now, in contrast with the above quasi-hyperbolic case. The direct meaning of this t^* here is that before this point in time the decision maker feels time passes quicker than it really does. Only one year has passed, but the decision maker feels it as 2.[1] Accordingly, the decision maker will expect higher compensation than what the exponential model would suggest, his/her discount factor will be lower, which comes to the same thing as in the quasi-hyperbolic case: before t^* the decision maker will be undercompensated, and after that overcompensated.

t^* will depend on the parameters of the model, g, h and r.[2] Although we cannot give an analytical solution for t^* in this case, t^* will definitely be decreasing in r, like in the quasi-hyperbolic case, and increasing in h, while decreasing in g, implying that these parameters might have the same function as β and δ in the quasi-hyperbolic model.

Loewenstein–Prelec case.

[1] "After all, I tried for 3 years! Seems like 30" – Jesus says in the musical *Jesus Christ Superstar*.

[2] If such a t^* exists at all. At certain combinations of the parameters g and h it is possible that $\alpha(t)$ and t will only have one common point, $t = 0$. In this case the decision maker is either constantly overcompensated independent of the time (if $\alpha(t)$ is always below t), or constantly undercompensated. It can be shown that this latter case is not possible: no such g and h combinations exist which would make a decision maker apprehend the progress of time constantly faster than it really is. The former case would mean the hyperbolic $\alpha(t)$ goes constantly under the exponential t in Figure 2. This is the case for example if $h < r$, there will be no intersection point except for the origin.

The main problem with hyperbolic models from my point of view is that papers related to these models only prove that experimental or field data analysed by them do not imply exponential discounting behaviour. Mostly the anomalies are explained by hyperbolic discounting, but the *g* and *h* parameters of the discount function are not estimated. The magnitude of these is undefined to such an extent that we can find a paper using $g = 4$ and $h = 1$ in a hyperbolic discount function (Angeletos et al. 2001), while another one uses $g = 10^5$ and $h = 5 \cdot 10^3$ (Laibson 1996).

Recapitulating, we can see that dynamically inconsistent time preferences can be grasped with models using discount rates depending on time, but can be formalised in various ways, and different authors give different functional forms to describe such discounting behaviour. I have shown that compensation offered by banks can be either too high or too low relative to the required compensation of the decision-maker, depending on the parameters of the discount function used in mental discounting. A direct consequence is that we will arrive at different results if we use such non-exponential discount functions as decision rules than if we assume the optimizing decision makers use exponential discounting. Behavioural economics conducted many experiments pointing out such anomalies: If the decision-maker discounted exponentially, he/she should have chosen to get sum *A* in time *x*, but he/she instead chose to receive sum *B* in time *y*. The same difference in mental discounting has to result in similar differences when evaluating flows of yields. In the next section I turn to the analysis of this problem.

2. Non-exponential discounting flows of income

The mainstream literature of hyperbolic discounting is concerned with evaluating and comparing single payoffs. Experiments conducted seek to determine the parameters of the discount function by examining what smaller but sooner receivable sum of money (*x* in the above terminology) is equivalent to what later receivable but larger sum (*X* in the above terminology) – these are called matching experiments –; or by examining which of a predetermined sooner-smaller or later-larger payoffs a decision maker would choose – these are called choice experiments (Read – Roelofsma 2003).

I am more interested now in the comparison of flows of incomes instead of single payoffs, according to the different discounting models.[1] How would an economic actor determine the present value of a regular flow of income (or any other kind of "utility flow") due for a certain period of time? In the first section I showed that if decision makers discount the future not according to the exponential but for example according to the quasi-hyperbolic model, then the present value of a single payoff calculated with the exponential model will be underestimated before time *t**, and overestimated thereafter. In this section I will show that this effect will compound and result in an increased distortion when calculating present value of flows of income, and that this distortion will consequently render the results of optimising behaviour according to the exponential model questionable.

---

[1] Although I will henceforth only consider the quasi-hyperbolic case; compared to the exponential, due to its above mentioned qualitative similarity to the hyperbolic discount functions, it is enough to show the arising problems for this case. Using different hyperbolic models would yield qualitatively similar results.

Let us take the simplest case as an example, that of a perpetuity. Here the conclusion is that the present value of a perpetuity discounted hyperbolically will be equal to its present value discounted exponentially if and only if

$$\sum_{t=1}^{\infty} C \cdot \delta_{exp}^{t} = \sum_{t=1}^{\infty} C \cdot \beta \delta_{hyp}^{t}$$

Determining the sums and dividing by the constant flow of income, then rearranging,

we have $\beta = \dfrac{r_{hyp}}{r_{exp}}$.

Since hyperbolic models assume that the long run interest rate is smaller than in the exponential case, and β is positive but smaller than one, combinations of the parameters can exist where the equation holds. In the case of combinations of parameters when it does not, that is, β is greater (smaller) than the ratio of the long run interest rates, the present value of the perpetuity calculated with the quasi-hyperbolic model will be greater (smaller) than calculated with the exponential model. Let us recall Figure 1! The above calculations mean we are calculating whether areas under the two depicted curves are equal.[1] The figure shows it may be possible, as before the point t* the quasi-hyperbolic curve is below the exponential, and afterwards above it. It is possible that the negative difference between the areas under the quasi-hyperbolic and exponential curves before $t*$ is exactly the same as the positive difference between them after $t*$. It is likewise possible that the negative or the positive difference is greater of the two. Obviously the discounting parameters determine which the case is.

The two areas under the appropriate curves vary with t, and thus we can determine the time $\tilde{T}$, as that point in time in the case of which the present value of an annuity from time 1 to $\tilde{T}$ would be equal calculated exponentially or quasi-hyperbolically. Since the exponential model underestimates the present values of payoffs before $t*$, this equality can only happen at a $\tilde{T} > t*$.

Likewise we can define $\vec{T}$ as the point in time in the case of which the present value of an annuity from time $(\vec{T} + 1)$ to infinity would be equal calculated with the two models. Again, since payoffs after $t*$ are overvalued with the exponential model, this equality will only hold at a $\vec{T} < t*$.

For this latter equality we can give a closed formula. The formula according to the above (and already simplified with the constant income flow) is:

$$\sum_{t=\vec{T}+1}^{\infty} \delta_{exp}^{t} = \sum_{t=\vec{T}+1}^{\infty} \beta \delta_{hyp}^{t} .$$

---

[1] Actually, since the quasi-hyperbolic case is non-continuous – as mentioned above – it would be more precise to speak of series instead of areas under a curve, and compare it with discrete-time version of exponential discounting. This is the reason why I used series instead of integrals in the calculation.

From which $\vec{T}$ is:

$$\vec{T} = \frac{\ln\left(\beta \cdot \dfrac{1-\delta_{exp}}{1-\delta_{hyp}}\right)}{\ln\delta_{exp} - \ln\delta_{hyp}} - 1.$$

Table 2 gives the values of $\vec{T}$ for some quasi-hyperbolic discounting parameters.

Table 2a: $\vec{T}$ values for $\delta_{exp} = 0.952$ (i.e.: $r_{exp} = 0.05$)

δ hyp \ β	0,25	0,5	0,75	0,8
0,96	9,36	-3,36	-10,81	-11,99
0,97	3,27	-7,42	-13,67	-14,67
0,99	-10,64	-18,77	-23,53	-24,28

Source: own computation

Table 2b: $\vec{T}$ values for $\delta_{exp} = 0.909$ (i.e.: $r_{exp} = 0.1$)

δ hyp \ β	0,25	0,5	0,75	0,8
0,96	151,53	64,32	13,31	5,19
0,97	49,50	11,64	-10,50	-14,02
0,99	-5,49	-23,39	-33,87	-35,53

Source: own computation

Negative values mean that in the quasi-hyperbolic case the present values of payoffs after point $t*$ are so much higher than in the exponential case that this could not be offset by lower present values before that: Quasi-hyperbolically calculated present value will always be higher.

3. A potential field of application: investment plans yielding external benefits

Until this point we saw the technical side of using non-exponential instead of exponential discounting. Next, let us turn to see to what theoretical problems this can be applied! First I show a simple numerical example to make my point. Let us imagine two mutually exclusive investment projects A and B, which would generate flows of income in the coming years for the investing firm according to Table 3.

Table 3: Cash flows (CF) of hypothetical investment projects A and B

year	0	1	2	3	4	5	6
CF (A)	-20	12	12	12	12	12	0
CF (B)	-40	14	14	14	14	14	14

Source: own computation

Let us now calculate the net present value for the two projects using an interest rate of 11% with the exponential discounting model! We find $NPV_A = 24.4 > NPV_B = 19.2$, which means that both projects have positive net present value but it is higher for

project *A*. If the projects are mutually exclusive and we rule out the possibility for now that two of project *B* can be made, the profit-maximising firm will chose project *A*.

Let us continue the above example and suppose that the projects will in further years generate external benefits to third parties. Utility of third parties can not be included in net present value calculations; however, they can form a substantial part of the overall welfare generated by any given project. Let project *A* generate 9 and project *B* 11 utilities, project *A* in the years 6 through 20 and project *B* in years 7 through 20. Let us incorporate these welfare yields using exponential discounting! As a result project *A* will still yield a higher overall welfare, 62.8, while project *B* only 60.3. If we, however, discount these external benefits for example with the quasi-hyperbolic method using β = 0.7 and an interest rate of 4.5%, that is δ = 0.957, we find that the overall welfare yield of project *A* is now 78.7 while for project *B* it is 79.7. In this case, project *B* can be considered to generate slightly higher overall welfare. *There exist values of the discounting parameters where the relation between exponentially calculated welfare and hyperbolically calculated welfare reverses.* This reversing is ultimately in connection with how the different yields are distributed relative to *t**, that is, whether the time horizon is right for the effects in connection with annuities mentioned in the previous section to play out.

The first important issue is to see whether this situation can arise for any plausible values of the discounting parameters. I consider it a further task to formalise the above problem analytically to find how robust the results are. The discounting parameters of the beta-delta discounting I used above, though, are comparable to those determined experimentally. Angeletos et al. (2001) use exactly these β and δ values, because "using annual periods, these parameter values roughly match experimentally measured discounting patterns" (p. 51). Another experiment (Laibson et al. 2007) finds surprisingly similar values for the parameters: "the MSM [Method of Simulated Moments] procedure yields an estimate of β = 0.703, with a standard error […] of 0.109. […] [for the other parameter one gets] an estimate of δ = 0.958, with a standard error of 0.007. The estimated values of β and δ imply a short-run discount rate of – ln(0.703*0.958) = 39.5% and a long-run discount rate of – ln(0.958) = 4.3%" (p. 17). The authors add about the result that "all of these estimates are statistically significant at the 1% level" (p. 4). Of course, both sources also examine what other factors can influence these parameters or how robust these estimates are, but use these values as benchmark, which means for us that it is at least not unreasonable to calculate with such values for β and δ.[1] It is certainly not sure whether these parameters are stable of if they can be aggregated to express the discounting behaviour of groups of people or society in general. I suspect that it varies with individuals and even decision situations.

A next important question is whether it is justifiable to use different discounting methods to monetary yields than utility yields, since money income and the profit of the firm will eventually buy utility. The reason why exponential discounting is used in cash flow analysis and the net present value theory is that banks use this kind of discounting and compound interest calculation when they gather or pay interest. Firms therefore have to be realistic and calculate with this model when assessing their future obligations or their opportunity costs. The exponential discounting method's interest rate is an

---

[1] For this combination of parameters we would find *t** = 6.931 and $\vec{T}$ = *-8.637.*

intertemporal rate of transformation for them enabling them to transfer future money to the present. However one plans to spread spending money over multiple time periods to ultimately yield utility, one is guaranteed to have maximum utility when following the net present value rule. In contrast, future utilities are by nature *not* transferable to present: one can not borrow utility against a prospective future income of utility. This is the reason I differentiate between firms discounting future incomes and consumers discounting future utilities, the former in an exponential way, the latter in a non-exponential way.

Recapitulating the section we can say that situations can exist, where a difference in the exponentially or non-exponentially discounted values of flows of income does matter. The remaining question is whether in real life situations exist, that bring up the same problem.

4. Non-exponential discounting and the government as regulator

In the above example whether we use exponential or non-exponential discounting, the profit-maximising firm will nonetheless choose project *A*, the one with higher net present value. Whereas firms in my view can not be expected to seek to maximise social welfare, much less to use non-exponential discounting, the government can assume this role. A government might find project *B* socially preferable over project *A* if the appropriate non-exponential discounting is applied. Thus the government might find it in its interest to regulate the firm or to try to persuade it to choose the project yielding higher overall social benefits while still remaining profitable. We have a problem similar to regulating a monopoly where social welfare is to be maximised under constraints on the monopoly's profit. Is such a scenario conceivable in connection with investment projects?

William D. Nordhaus in his 1967 paper "The optimal life of a patent" describes a situation exactly like this. In his case the government can set the time for which it grants a patent, that is, intellectual property protection to an inventor. Once the government sets this decision variable (how long a patent lives, for what length of time a firm can exclusively use and profit from its inventions), the firm will decide how much it should optimally spend on research and development. In the model the profit maximising firm will find that the longer it can enjoy monopoly rights in the usage of its invention, the more it is worth to spend on R&D. Longer patent life granted by the government results in higher R&D, which leads (in the model) to higher cost saving and more profit, which will also be appropriable for a longer period of time because the longer patent life initially set. Our example above mimics these properties. At the same time, even because of the monopolistic power of the inventor granted by the patent rights, the firm also creates deadweight loss, which can only transform into increased consumer surplus after the patent right expires. This increased consumer surplus is, however, external to the firm (just as it was treated in the above numerical example), and only enters the welfare calculations of the regulating government to decide the patent length. Thus the circle closes: The longer the life of the patent, the more innovation will result, which will generate ever higher potential future consumer surplus gains to the consumers, but they also have to wait ever longer to acquire this surplus because of the longer patent life.

The firm will decide about resource allocation to R&D based on expected profits, whereas the government also takes into account the discounted value of future prospective increased consumer surplus when setting the life of the patent to optimal. If the government not only takes this into account, but calculates it as the consumers themselves (mentally) would, for example quasi-hyperbolically, then its decision about the length of patent life could change. For any given level of R&D the optimal life of a patent grant can be different when the associated welfare gains are taken into account differentially discounted. The government as a regulator can be able to internalise the future welfare effects more appropriately for the firm.

The above fictitious numerical example simplifies the problem to the extremes, but in my opinion the theory of optimal patents is an appropriate field to examine how the usage of non-exponential discounting models modifies economic policy. Nordhaus's model formalises all necessary relations between patent length and the firm's decision as to how much resources to invest in R&D, what returns to the firm it will generate, how this return relates to further external benefits to the consumers. Author's firm determination is to incorporate non-exponential discounting into this basic model to come up with a more formalised statement on how much difference discounting non-exponentially in an already well-established model would make.

Summary

The aim of my study was to incorporate a behavioural economic observation into an economic application. Behavioural economics states that in experiments studying time preferences subjects repeatedly make decisions that do not match forecasts based on the conventional assumptions. When choosing among a sooner-smaller and later-larger reward the data fit better if we make different assumptions about discounting behaviour.

The starting point of my study was to see how theoretical results of intertemporal choice are modified if we take these different discounting behavioural assumptions as given. To see this I used hyperbolic and quasi-hyperbolic discount functions to calculate present values of annuities and perpetuities. I then showed how this present value varies depending on the different discounting parameters and that in turn the exponentially discounted present value is either higher or lower than the non-exponentially discounted.

Transferring the obtained annuity present value formulae to investment project evaluation, I showed that if utilities or costs associated with an investment project are accrued during a longer time, and it is reasonable to discount these (or part of these) non-exponentially, then this can influence the net present value and maybe even the investment decision itself. I presented a fictitious investment decision situation where in the evaluation of two projects the exponential discounting favoured the one but the non-exponential discounting the other investment project.

Finally I wanted to point out that such a situation is not only a play with numbers, but can really occur: the logic of setting the optimal patent life is identical.

My concluding qualitative statement was that economic policy-makers can encounter situations when they are to be the actors that take into account the consequences of non-exponential discounting, and also the actors that can possibly incorporate these external effects and make it count for the other economic actors too.

References:

ANGELETOS, GEORGE-MARIOS, LAIBSON, D., REPETTO, A., TOBACMAN, J., WEINBERG, S. (2001). The Hyperbolic Consumption Model: Calibration, Simulation and Empirical Evaluation. *Journal of Economic Perspectives*. 15 (3), pp. 47-69.

CAIRNS, J. (2006). Developments in discounting: with special reference to future health events. *Resource and Energy Economics*. 28 (3), pp. 282- 297.

FREDERICK, S., LOEWENSTEIN, G., O'DONOGHUE, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*. XL (June), pp. 351-401.

HARVEY, C. M. (1995). Proportional discounting of future costs and benefits. *Mathematics of Operations Research*. 20 (2), pp. 381–399.

KIRBY, K., HERRNSTEIN R. J. (1995). Preference Reversals Due to Myopic Discounting of Delayed Reward. *Psychological Science*. 6 (2), pp. 83-89.

LAIBSON, D. (1996). *Hyperbolic Discount Functions, Undersaving and Saving Policy*. NBER working paper No. 5635.

LAIBSON, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*. 112 (2), pp. 443-477.

LAIBSON, D., REPETTO, A., TOBACMAN, J. (2007). *Estimating Discount Functions with Consumption Choices over the Lifecycle*. NBER working paper No. 13314

LOEWENSTEIN, G. F., PRELEC, D. (1992): Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*. 107 (2), pp. 573-597.

MAZUR, J. E. (1987). An adjusting procedure for studying delayed reinforcement in M. L. Commons, Mazur, J. E., Nevin, J. A., Rachlin, H. (eds.): *Quantitative Analyses of Behavior. V. The Effects of Delay and of Intervening Events on Reinforcement Value*. Erlbaum, Hillsdale NJ, pp. 55–73.

NORDHAUS, W. D. (1967). *The Optimal Life of a Patent*. Cowles Foundation Discussion Papers 241. New Haven.

PHELPS, E. S., POLLAK, R. A. (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies*. 35, pp. 185-199.

RACHLIN, H. (2006). Notes on discounting. *Journal of the Experimental Analysis of Behaviour*. 85 (3), pp 425-435.

READ, D. (2001). Is Time-Discounting Hyperbolic or Subadditive? *Journal of Risk and Uncertainty*. 23 (1), pp. 5-32.

READ, D., ROELOFSMA, P. H. M. P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes*. 91, pp. 140-153.

SAMUELSON, P. A. (1937). A Note on Measurement of Utility. *The Review of Economic Studies*. 4, pp. 155-161.

THALER, R. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters*. 8 (3), pp. 201-207.

TROPE, Y., LIBERMAN, N. (2003). Temporal Construal Theory of Time-Dependent Preferences. In: Brocas, I., Carrillo, J. D. (eds): *The Psychology of Economic Decisions. Volume I*. Oxford: OUP. pp 235-249.

HYPERBOLIC DISCOUNTING AND ECONOMIC POLICY

Benedek Nagy

University of Szeged, Hungary, 6722 Szeged, Kálvária Av. 1., Nagy.Benedek@eco.u-szeged.hu

Abstract: Economic policy-making often entails comparison between immediate costs and flows of future benefits or immediate benefits and series of future costs. Economics has a tool to handle such comparisons: the present- and future value calculations and the net present value rule. Experimental economics, however, has strongly criticised the method of exponential discounting applied in such calculations. Based on experiments for the sake of more psychological realism, they propose alternative methods to the exponential model: hyperbolic and quasi-hyperbolic discounting models.

The present paper has a twofold objective: first, to review these different models and the relationships between them to show how the different models will yield different results when calculating and comparing present values of a single future payment, but even more if we compare present values of flows of future payments. The literature has not yet employed the hyperbolic and quasi-hyperbolic models for such calculations. Second, I point out why it is important to heed the findings of experimental economics especially in the field of economic policy-making.