Forthcoming in D. Pannell and S. Schilizzi, eds. (2005/6), *Discounting and Discount Rates in Theory and Practice*. Cheltenham: Edward Elgar.

6. Reconsidering reconsidered: Why sustainable discounting need not be inconsistent over time

John C.V. Pezzey

SUMMARY

Economists assess dollar costs and benefits in the future by discounting them. Each cost or benefit is multiplied by a discount factor appropriate to the time when it occurs, and then total discounted costs are subtracted from total discounted benefits to get net present value. The rate by which the discount factor declines over time is the discount rate. Economists also apply discounting to costs and benefits measured in terms of wellbeing, and the discount rate is then called (the rate of) impatience.

Whether or not impatience should be constant over time is important to long-run development theory. If non-renewable resources are essential inputs to production, then the return on capital investment will eventually decline towards zero, and thus fall below any constant rate of impatience. Once this happens, and assuming the economy always maximises net present value, then wellbeing will be falling over time. So there is considerable interest in 'sustainable discounting': using a falling rate of impatience that stays below the return on capital investment, and thus gives a sustainable development of, rather than an eventual fall in, wellbeing.

However, any non-constant rate of impatience typically causes inconsistency, where people make a plan now for their entire future, but on reconsidering it later, want to change their plan. I show diagrammatically that inconsistency occurs only if people use a relative time approach to reconsidering, by using a discount factor that depends only on how far off future costs and benefits are relative to the moment of reconsidering. If instead people use an absolute time approach, where the discount factor

depends only on the date of some future cost or benefit, then no inconsistency occurs. So sustainable discounting using the absolute time approach offers a way of achieving sustainable development where conventional, constant discounting would make it impossible. However, there is little evidence so far that people are prepared to constrain their impatience as may be required by sustainable discounting.

6.1 INTRODUCTION

It is well known that using a constant discount rate in cost-benefit analysis is the bane of taking the very long-term future seriously in economic decisions. As long as the future is far enough away, such ‘constant discounting’ results in even a huge future value (cost or benefit) having only a tiny present value. This tiny value gets added to amounts incurred today and at all other times to reach a net present value, which we assume people try to maximize when making plans for the future. Moreover, under constant discounting, an amount’s present value keeps on shrinking by a multiple of typically more than 10 for each century that the amount’s date recedes into future, even if that century is already a thousand years from now. (A modest 2.5 per cent discount rate per year results in a 91 per cent discount rate per century.)[1]

Many people are instinctively uneasy about constant discounting. This is especially true when it is applied to the distant future, where it makes such a difference; and also when it relates to wellbeing (utility, in economists’ jargon) rather than real dollars, an application first proposed by Samuelson (1937) and commonly used by economists since. Discounting utility is often called pure time preference, or simply impatience, referring to the fact that most people do prefer a unit of wellbeing (utility) now rather than later. A powerful reason for unease with constant utility discounting (i.e., a constant rate of impatience) comes from two standard results in the theory of any market economy that grows so as to maximize the total present value of people’s utility.

6.2 CONSTANT VERSUS DECLINING DISCOUNT RATES: THEORY AND EVIDENCE

The first result states, subject to some simplifying conditions, that the growth rate of consumption is proportional to the (rate of) return on capital investment minus the rate of impatience[2]. The second result is that if some kind of non-renewable resource is an essential input to the functioning of the world economy, then there is reason to expect that the return on capital will eventually decline towards zero in the long term (Dasgupta and Heal, 1974). (By non-renewable resources, one nowadays thinks of the world’s remaining capacity to absorb greenhouse gas emissions without a catastrophic change in climate, or the amount of biodiversity not yet destroyed by land clearing and conversion, as much as stocks of petroleum or platinum ore.) If these two theoretical results hold true and the rate of impatience is constant, then sooner or later this rate must outweigh the declining return to capital, and consumption will then decline forever. This means that 'later generations (should they exist) suffer incredibly as a result of the initial profligacy under the [constant discounting] programme' (Dasgupta and Heal, 1979, p. 299).

In fact, as Weitzman (1998, p.1592) points out, there is little evidence so far of any such long-term decline in the real rate of return, at least for the USA over the past century or so. But that does not mean it will not decline in the far future. If it does decline, a possible solution to the bleak, long-term future that the world would otherwise face because of a constant impatience rate, would be for the rate of impatience also to decline, and fast enough to stay below the declining return on capital. For a long time there was little factual evidence as to whether typical impatience rates are constant or declining. Over recent years, however, there has been growing empirical (real-world) and experimental (laboratory) evidence that, in many situations, a person's rate of impatience does decline over time, at least when she or he contemplates the remainder of their own lifetime from where they are now (Frederick et al., 2002, Section 4). That is, the discount rate that most people apply when comparing joys and woes expected (say) 50 rather than 40 years from now is generally lower than the rate they apply when comparing joys and woes expected 20 rather than 10 years from now.

However, a crucial theoretical problem then arises. Strotz (1955/6) showed that a non-constant rate of impatience typically makes people's choices over time inconsistent. They may make a plan now for their entire future, but regret it later, and want to change their mind when some point of the future arrives. Frederick et al. gave ample evidence that such time-inconsistency does occur, and that it results in people using binding contracts and other pre-commitment devices to stop them forever undoing their past decisions.

6.3 ABSOLUTE VERSUS RELATIVE DISCOUNTING

My purpose in writing this note is to highlight another solution to the supposed problem of time inconsistency caused by non-constant discounting. My explanation focuses on a simple graph, Figure 6.1 below[3]. The consistency or otherwise of decisions made using a non-constant rate of impatience depends on what kind of rule people use when reconsidering an earlier decision at a later date. Figure 6.1 assumes that the average person's discount factor, which they multiply by their measure of wellbeing at any moment to calculate their present value of utility, then falls continuously over time (marked t on the horizontal axis) along curve ABCD. The discount factor on this graph is not constant: it falls from 1 initially, to ½ at time $t = 1$, ⅓ at time $t = 2$, ¼ at time $t = 3$, and $1/(1+t)$ at any other time t. The shape of the $1/(1+t)$ curve is a type of hyperbola, so the use of this discount factor is one example of what is called 'hyperbolic discounting'. The discount rate per time period is then also falling (i.e., non-constant), from 1 – (½/1) = 50 per cent between times 0 and 1, to 1 – (⅓/½) = 33⅓ per cent between times 1 and 2, to 1 – (¼/⅓) = 25 per cent between times 2 and 3, and so on. (If instead the discount factor declined exponentially over time, from 1 to ½ to ¼ to ⅛ in successive time periods, then the discount rate would be constant at 50 per cent over every time period.)

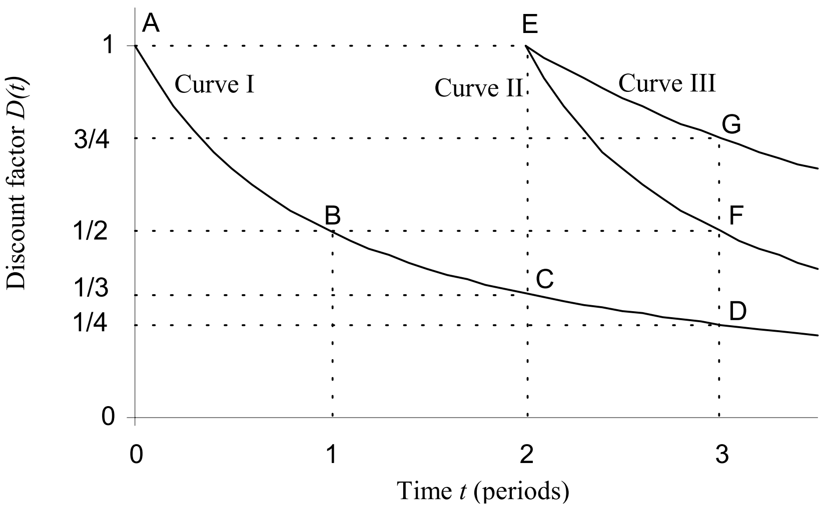

Figure 6.1 Conventional and alternative approaches to discount factors used to reconsider future plans

LEGEND

Curve I: $D_0(t) = 1/(1+t)$, the original discount factor curve used when making future plans at $t = 0$

Curve II: $D_0(t-2) = 1/(1+t-2)$, the conventional, relative time discount factor curve used when making plans at $t = 2$

Curve III: $D_0(t)/D_0(2) = (1+2)/(1+t)$, the alternative, absolute time discount factor curve used when making plans at $t = 2$

The crucial issue with non-constant discounting is this. If the person reconsiders her future plans, say after two time periods at t = 2, what discount factor curve will she use then? The conventional answer, which I will call the relative time approach, is that she would restart her discounting clock by using discount factor curve EF, which plots the value of $1/(1+t–2)$ from t = 2 onwards. EF is the old discount curve ABCD shifted two time periods to the right. It starts at a value of 1, so that there is no discounting of amounts at t = 2, the new 'now' or starting time, and the discount factor applied 'now' to an amount at some future time then depends only on how far away that time is relative to 'now'. An alternative, absolute time approach, would be to use curve EG, which plots the value of $(1+2)/(1+t)$ from t = 2 onwards. EG also starts from 1, and is the curve CD originally used to compare events between t = 2 and t = 3 when making plans at t = 0, expanded upwards by a factor of 3. The discount factor applied 'now' to some time in the future then depends not only on how far away the time is relative to 'now' (t = 2), but also on how far away it is from the absolute origin of time (t = 0).[4]

The effect of relative versus absolute discounting on time consistency can be readily seen from the following example. Under the relative time approach (curve EF), when reconsidering plans at t = 2, the new present value ratio of an amount at t = 3 compared to the same amount at t = 2 is ½/1 or 50 per cent, much lower than the ¼/⅓ or 75 per cent originally used at t = 0. So the new plan will indeed be inconsistent with the original plan. By contrast, under the absolute time approach (curve EG), the ratio of present values between t = 3 and t = 2 is ¾/1, also 75 per cent. So the new plan for t = 2 will be consistent with the original.

6.4 SUSTAINABLE DISCOUNTING: REALITY CHECK

The above result can be summarized as follows: the absolute approach to discounting makes non-constant discounting time-consistent, and may thus resolve an important moral dilemma in economies where accumulating capital and knowledge can substitute indefinitely in production for dwindling inputs of a non-renewable natural resource. Many people (but not all, depending on views about long-run resource scarcity) see such substitution as vital to sustaining modern economies in the far future. The dilemma for such economies is that on the one hand, an objective of constant consumption forever can be achieved under non-constant discounting, by a high enough level of saving, and subsequent investment in capital and knowledge. However, constant consumption is often seen as perpetuating poverty or giving foolishly conservative injunctions, by preventing the economy from ever growing (Solow, 1974). On the other hand is the inevitability of the Dasgupta-Heal result from constant discounting mentioned above: that consumption will ultimately dwindle to zero, which is manifestly unfair to future generations when constant consumption is feasible. Pezzey (2004) gave an example of how this dilemma can be avoided by using absolute hyperbolic discounting at a slow enough overall rate, resulting in steady, sustainable consumption growth, without any problems of time-inconsistency[5].

But in everyday language, for pure time discounting to depend on absolute time would mean that the average person's impatience would depend not just on their own psychological preferences, as influenced by their age, health, wealth and so on, but also on absolute time and hence on how the sustainable the world is then. If the long-run future of the world is indeed seen as broadly that of a capital-resource economy, then the main problem for intergenerational equity is the Dasgupta-Heal result that the return to capital investment will eventually decline as capital is accumulated and non-renewable resources are gradually used up. An absolute approach to discounting – sustainable discounting, if you will – would require people to constrain their impatience within what the world can sustain at the time when they happen to be living. Such a constraint on discounting would arguably lie at the heart of a true sustainability ethic.

However, despite the almost universal proclamation nowadays of sustainability as a paramount policy goal, there is little evidence yet that any kind of constrained, absolute kind of impatience is emerging in any real economies. It is notable that in their authoritative fifty-page review of the huge complexities of time discounting and time preference, Frederick et al. (2002) focused solely on psychological determinants of impatience, and made no mention of any connection between the state of the world and discount rates. This concurs with my own casual observation over two waves of global environmental concern, in the late 1960s and late 1980s, to date. I see few signs that the average consumer or voter in 2005 discounts the next five or fifty years any less than she did in 1965 (think only of the recent politics of petrol price rises). But one can always hope that the theoretical possibility shown above – that of consistently discounting the far future at a rate no larger than the world can sustainably afford – may eventually become true in reality!

ACKNOWLEDGMENT

I thank Dave Pannell, Steve Schilizzi and Cees Withagen for helpful comments. The usual disclaimer applies.

NOTES

1. The terms here have precise mathematical meanings. An amount (of real dollars, or well-being) occurring at time *t*, say *A*(*t*), is multiplied by a *discount factor*, say *D*(*t*), to give *A*(*t*)*D*(*t*), its *present value* at time 0. The *discount rate* is $-\dot{D}(t)/D(t)$, the rate at which the discount factor declines at time *t*. Using a negative exponential discount factor, $D(t) = e^{-rt}$, with *r* being a positive constant, is called either 'constant discounting' or 'exponential discounting', and *r* is then the constant discount rate.
2. Economists will recognise this word equation as the Ramsey rule or Euler equation for the optimal development of a present-value-maximizing economy (e.g., see Equation (2.8) in Barro and Sala-i-Martin, 1995), where for simplicity any direct (amenity) influence of the environment on well-being is left out.
3. I give a formal mathematical treatment in Pezzey (2004).
4. Mathematically, if $D(x,y)$ is the discount factor applied at time *x* to an amount occurring at future time *y* (> *x*), then Strotz's requirement is that $D(.)$ should be a function only of (*y* - *x*), the distance that time *y* is in the future relative to time *x*, but not of *x* and *y* independently. An equivalent requirement is the stationarity axiom in Koopmans (1960), which says that the ranking of two futures should not change according to how far away they both are. Koopmans showed that constant discounting is the only type of discounting consistent with stationarity and several other axioms. Using relative time discounting on Figure 1 breaks stationarity. For example, a future with (6, 36) utils at times *t* = (1, 2) respectively has a present value, considered at *t* = 0, of 6/2 + 36/3 = 15 utils. It is thus preferred to a future with (16, 18) utils since the latter has a present value of 16/2 + 18/3 = 14. But when reconsidered at *t* = 1, these two futures have present values of 6 + 36/2 = 24 and 16 + 18/2 = 25 utils, so now the second one is preferred. But as argued in Pezzey (1997) and Frederick et al. (2002), an axiom is only an axiom, and it may not be empirically valid when tested against data: neither Samuelson nor Koopmans claimed any empirical validity for their elegant theoretical results on constant discounting.
5. Withagen et al. (2003) showed that Solow's (maximum) constant consumption path can be the outcome of present value maximization with a decreasing discount rate. For specific functional forms, Pezzey (2004) extends this result to growing consumption.

REFERENCES

Barro, R.J. and X. Sala-i-Martin (1995), Economic Growth, New York: McGraw-Hill.

Dasgupta, P.S. and G.M. Heal (1974), 'The optimal depletion of exhaustible resources', Review of Economic Studies, 41, Symposium on the Economics of Exhaustible Resources, 3-28.

Dasgupta, P.S. and G.M. Heal (1979), Economic Theory and Exhaustible Resources, Cambridge: Cambridge University Press.

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

Koopmans, T.C. (1960), 'Stationary ordinal utility and impatience', Econometrica, 28, 287-309.

Pezzey, J.C.V. (1997), 'Sustainability constraints versus 'optimality' versus intertemporal concern, and axioms versus data', Land Economics, 73, 448-466.

Pezzey, J.C.V. (2004), 'Exact measures of income in a hyperbolic economy', Environment and Development Economics, 9, 473-484.

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

Solow, R.M. (1974),'Intergenerational equity and exhaustible resources ', Review of Economic Studies, Symposium on the Economics of Exhaustible Resources, 41, 29 –46.

Strotz, R.H. (1955/6), 'Myopia and inconsistency in dynamic utility maximization', Review of Economic Studies, 23, 165-180.

Weitzman, M.L. (1998), 'On the welfare significance of national product under interest-rate uncertainty', European Economic Review, 42, 1581-1594.

Withagen, C., G.B. Asheim and W. Buchholz (2003), 'On the sustainable program in Solow's model', Natural Resource Modeling, 16, 219-231.