Preference cyclicity and money pumping in a virtual economy

David Spurrett*(1), John McCoy*, Don Ross** and Ben Murrell*

* School of Philosophy and Ethics, University of KwaZulu-Natal, ** Departments of Philosophy and Economics and Center for Ethics and Values in the Sciences, University of Alabama at Birmingham and School of Economics, University of Cape Town. (1) Corresponding author: spurrett@ukzn.ac.za

Abstract

Agents with transitive or cyclical preferences will pay to undo exchanges for which they had previously paid transaction costs. Such agents can be money pumped actively, by a knowing party that offers them a sequence of exchanges tailored eventually to strip them of all their assets, or passively by the accumulation of cases in which they reverse previous exchanges at some cost, as long as some of the agents with which they trade have acyclical preferences. We describe preliminary modeling work with a virtual economy in which agents with different utility functions forage for goods of different types, and also engage in exchanges. Some of the agents discount delays, both in the case of foraging decisions and in evaluating exchanges, hyperbolically, others exponentially. Using this model we pose a number of questions regarding the conditions under which agents with cyclical preferences due to hyperbolic discounting are significantly disadvantaged, and whether those conditions confer compensating advantages. We identify ranges of foraging problems in which hyperbolically discounting agents tend to outperform exponentially discounting ones. This suggests a basis for evolutionary explanation (both biological and institutional, in different instances) of widespread empirical observation of hyperbolic discounting.

(1) Introduction

This paper is not a work of economics, although it is occasioned by on the one hand a piece of economic theory (money pump arguments indicating that maladaptiveness of preference cyclicity) and on the other a piece of behavioral economics (empirical work on discount functions). Neither is it quite the paper we envisaged at the time the original abstract was drawn up, since the work we then envisaged required the model we develop and use to be fully functional earlier than it eventually was. That paper would have been a fragment of the artificial life literature inspired by some specifically economic problems. This paper, instead, is a largely narrative account of the early stages of a research programme. We hope that the fact that research of the sort in question is relatively rare in South African economics makes a report of this nature worth some attention.

Economists don't need telling that preference cyclicity is a bad thing. They're all taught that agents with cyclical preferences are vulnerable to money-pumping, and that such agents should therefore be driven out of markets including the ongoing high stakes market that is the tournament of evolution. Indeed, there is a clear sense in which from the perspective of mainstream economic theory, a putative agent without stable preferences isn't really an agent at all, since it would defeat applications of RPT.

The core of the traditional money-pump argument (See Ross 2005, Chapter 5) goes as follows:

- Let *X* be an agent that prefers *a* to *b*, prefers *b* to *c* and prefers *c* to *a* (where *a*, *b*, and *c* are bundles). That is, let *X* have cyclical preferences, and some initial stock of *a*, *b*, and *c*.
- Suppose another agent (with acyclical preferences)[1] offers *X* a series of trades, where *X* first pays to exchange stock of *a* for *b*, then of *b* for *c* and then finally of *c* for *a*. Then at the end of the sequence *X* has less of everything than to begin with.
- This makes *X* a 'money pump', such that repeating the sequence would eventually drain *X* of all, or all except an infinitesimal fraction, of its initial asset stock (assuming that such an agent could have had any stock to start with).

Agents with acyclical preferences, that is, will be favoured over ones with cyclical preferences. We are generally supposed to conclude from this that the latter should go extinct (be pushed out of markets, or excluded by natural selection) in the event that the former are possible at all.

Given this argument as background, the empirical study of discounting (especially for delays) provides substantial confirmation for a very perplexing result.

Since waiting, or more generally the passage of time, involves accumulating uncertainty, future rewards should be valued less in the present. Individual variation in risk-sensitivity can make curves drawn through measures of the utility associated with identical goods at different times decline with varying steepness, but as long as each curve declines exponentially[2] no two curves

---

[1] It is, of course, not essential that what works the pump itself be a single agent, only that a series of trades of the required form takes place. So, for example, three agents such that one prefers *b* to *a*, the next *c* to *b* and the third *a* to *c* could take turns trading and thereby amount to a pump of *X* while not themselves being vulnerable.

[2] A standard discount function is called 'exponential' because future value is determined by application of a linear power function to undiscounted (present, or zero delay) value. For example:

$$V = A(1 - k)^D$$

will ever cross. Curve crossing when curves represent value as a function of delay would amount to reversing preference purely as a function of the passage of time, and hence to a failure of preferency acyclicity.

Evidence shows that people and various other animals including rats, various kinds of monkey and pigeons, do, however, discount non-exponentially. The experimental literature suggests, more specifically, that the curve best describing the data is hyperbola-like[3] (Ainslie 1992, Myerson & Green 2003). For simplicity of exposition we'll call the 'real' discount functions hyperbolic. The experiments in question involve behavioural work with nonhuman animals (e.g. Mazur, 1997) emitting behaviours in response to controlled reward schedules, and a variety of work with humans involving procedures designed to locate set of indifference points for pairs of amounts and delays (Green *et al.*, 1994; Kirby & Herrnstein, 1995; Kirby & Marakovic 1995; Mazur, 1987) to which sets of points curves can then be fit. The curve is regarded as a delay-specific variation of the Herrnstein 'matching law' (Chung & Herrnstein, 1967), stating that behaviour is allocated between alternatives in proportion to reward received from them.

Hyperbolic delay discount curves for two rewards (depending on the relative undiscounted magnitudes and positions in time of the two rewards) can cross, even for rewards in the same modality. That is, the passage of time alone can lead to preference reversal. The predicted reversal has itself been observed in animals (Ainslie & Herrnstein, 1981) and in human choices for various types of rewards including fruit juice, access to computer games, and both real and hypothetical money (Forzano & Logue, 1992; Millar & Navarick 1984; Ainslie & Haendel, 1983, see also Ainslie 2001).

In the light of the money pump argument outlined above, this propensity for preference reversals should be regarded as maladaptive. What is going on?

There have, unsurprisingly, been various suggestions to the effect that there may be considerations that weigh in favour of hyperbolic discounting, offered variously to account for experimental data, to adaptively explain them, or otherwise to suggest a response to the money pump argument. Among the suggestions in the literature (this is by no means a thorough survey) are the following:

- On some accounts hyperbolic discounting is to be explained by reference to the increasing costs (in risk) of waiting (Kagel et al., 1986; Logue, 1988). Recent neuroeconomics suggests that a nearly hyperbolic discount function could be composed out of the interaction of an exponential function for 'mere' delay, and a diffusion function for decreasing probability of delivery with increasing delay (Montague and Berns 2002).
- Stephens and collaborators (Stephens & Anderson 2003; Stephens & McLinn 2003, Stephens et al, 2004) suggest that in a foraging situation with patches of varying quality and 'stay or go' choices that impatience may be adaptive, and have found in various experiments with captive blue jays that birds who preferred immediate rewards (for

---

In this formula V is the value at the later time, A the undiscounted value (at zero delay) k a discount rate, and D the delay.

[3] Mazur (2001) provides a simple (hyperbolic) formula fitting the experimental data well:

$$V = \frac{A}{1 + kD}$$

In this formula V is value at no delay, D is the delay, and k is a constant representing degree of impatience. For a review of empirical results and the use of different formula to account for the results, see Green & Myerson (2003).

several operational versions of ‘immediate’) did at least as well, or better, than ones that waited for larger later rewards.

- Stevens et al (2005) found that two species of primate (tamarins and marmosets) had different discount rates. Since the two species share many behavioural, ecological and phylogenetic features, this is perplexing, but Stevens et al argue that the fact that tamarins mostly eat insects (requiring quick action) while marmosets eat sap that flows slowly from trees, may indicate that the greater preference for immediacy by tamarins makes adaptive sense.
- Ainslie & Herrnstein (1981) suggest that low discount rates would be maladaptive because of the increased information processing costs of tracking reinforcers at greater and greater distances (in time and space), suggesting that the discounting steepness that is adaptive for a given agent will be a function of the rate of change in its environment, and propose further considerations that might motivate a hyperbolic curve, rather than simply a steep exponential one.
- Ainslie (2001) proposes that generally hyperbolically discounting agents would be more likely to act in ways that meant risks for them as whole organisms (for example in desperate defence of their young, or in taking a nearby opportunity for copulation), but which tereby favoured the interests of their genes better than otherwise.[4]

Not all of these proposals amount to arguments for *hyperbolic* discounting specifically, rather than for steeply exponential discounting. Not all of them have been quantitatively specified or subject to empirical testing tested.

The work we report on here is an attempt to begin to pursue *some* of the questions concerning the relations between money pumping arguments and the ecological consequences of hyperbolic discounting.

(2) Our project

Our chosen research instrument is the spatially explicit agent based model (see Axelrod forthcoming). Our overall aim was to make a start in constructing an artificial world in which a population of agents differing in their discount functions faced both a foraging problem and had opportunities to engage in exchange with one another, in order to observe the effects of various manipulations of the foraging environment, the peoperties of the agents themselves, and the mechanisms of trading, on the net effectiveness of the agents depending on their discount functions. We hoped that by doing this we would eventually be able to shed some light on the credibility of some of the hypotheses supposedly favouring hyperbolic discounting above, especially those that had not yet been subject to any sort of quantitative examination.

Our more specific initial aim was to construct a world populated by agents with differing discounting dispositions in order to ask very simply what difference these dispositions in fact made to their success.

The present project is a small part of a larger project to construct a working model of the inter-temporal dynamics of a ‘picoeconomic’ agent, as sketched in Ainslie (1992, 2001). Ainslie’s proposal starts from the fact of hyperbolic discounting, and suggests that human agents are best understood as consisting in populations of competing interests. While of considerable interest, his work has not been framed in a form suitable for modelling or other quantitative evaluation, and we hope ultimately to help remedy this shortfall.

---

[4] The discovery by Wilson & Daly (2003) that priming with a task involving looking at pictures of attractive women leads to temporary increase in discounting steepness in males lends some weight to this hypothesis.

(3) Description of the model

Our hugely simplified model has been constructed in *NetLogo* (Wilensky 1999). *NetLogo* is a purpose built agent-based modelling language and environment, intended for use in a modelling a range of phenomena, especially complex systems with large numbers of interacting parts. A standard *NetLogo* model consists of a world (a 100 by 100 grid of patches) and a set of agents (generically known as 'turtles') and a suite of procedures specifying how the patches and turtles act and interact.

In our world most patches are empty, but specified numbers of patches contained either 'large' or 'small' quantities of two different sorts of reward, which we named apples and oranges. (Hyperbolic discounters will prefer some small rewards available soon to larger ones available later when the small reward is immanent, but not at other distances. The presence of two different patch sizes in the model was an attempt to make these propensities of hyperbolic discounters active, so as to explore what their effects were.) These food patches are the resource collected by the turtles.

The population of turtles is divided into two groups, hyperbolically discounting 'hedgehogs' and exponentially discounting 'elephants'. The population also divided into groups, each with a characteristic indifference point between apples and oranges. This gives a total of four 'breeds' of turtle: *hhogs1, hhogs2, elephants1, elephants2*. So, *hhogs1* and *elephants1* are indifferent between a_1 apples and b_1 oranges, and *hhogs2* and *elephants2* are indifferent between a_2 apples and b_2 oranges and so.

The turtles gather food patches by moving towards whatever patch is most attractive at any time, and 'eating' any food patch they happen to be standing on. Any patch that is eaten is immediately replaced by another with the same flavour and magnitude in a random (but unoccupied) location. This is to allow arbitrarily long runs of the model, and also to dilute the effects of contingencies about the initial random distribution of turtles and patches from being regarded as results about the foraging dispositions of the turtles.

The turtles are also able to offer and accept trades with each other, once they have gathered some food. Finally, the model can run in a 'no foraging' condition, where all turtles are given an initial handout and trade without foraging, and in a 'no trading' condition where foraging takes place without trading. These options serve a primarily diagnostic function in testing individual sub-systems in the simulation. In a little more detail the model works as follows:

Foraging

The foraging agents can 'see' all the patches in the environment, but not the other agents. A foraging agent calculates the discounted value (given its particular discount function, and impatience) to it of all the patches in the environment taking into account its preferences in terms of apples and oranges and orients towards the patch which has the highest discounted value for it.

At each timestep (each increment of the counter) agents both move forward 1 unit of distance, taking them closer to what at that time is the most preferred patch. Since all agents move at the same speed, the calculations of discounted value can safely use distance as the measure of delay. The valuation process is repeated prior to every forward step, to allow for the possibility that someone else has eaten the patch, or a new one arisen nearby, hence leading to revised priorities given the same preferences, and to allow for revised preference rankings in agents with non-exponential discounting, for whom the changing relative distance of patches may have precipitated preference reversals.

Trading

The process of trading in the model is handled by constructing and operating on a series of nested lists.

First, each turtle generates a 6-tuple list (id, flavour_o, amt_o, time_o, flavour_w, amt_w, time_w) where flavour_o is the flavour (apples or oranges) the agent 'wants' to offer, amt_o is the amount (of flavour_o) it wants to offer, and time_o the length of time from now that the agent wants to hand over what is offered. Similarly flavour_w is the flavour that the agent wants, amt_w the amount and time_w is the time, from now that it wants to receive the specified goods. This list is then composed into a list comprising all of the 'offers' in a given round.

This list is constructed semi-randomly, but the selection of random values is constrained in several ways:

- First, it is constrained by the discount functions and apple-orange indifference of the agent in question. No agent will generate an initial list describing an exchange that it would not prefer to make than not to make given the delay from the present to the specified events.
- Second, it is constrained by the 'budget' of the turtle, in a simple crude way – no turtle can offer to pay something at any time that it does not have on hand at the time of the deal.[5]
- Third, it is constrained by parameters in the model (that can be manipulated during a simulation) that specify how 'greedy' the turtles are, which is to say how much better than indifference (as a percentage) a trade has to be before they will take it, and the maximum distance in the future it is possible for part of a trade to take place.

Thereafter the question of what deals if any to accept are handled by allowing each turtle to inspect the list of deals, and calculate whether it would prefer to accept or reject it, accepting the one that it prefers the most (if any are preferable at all in that round). The order in which the turtles inspect the deals list is randomised from round to round, in order to eliminate systematic favouring or disadvantaging of any turtles as a result of regularly getting an earlier look at the list of options.

Accepted deals are then transformed into a list of pairs (i.e. two for each accepted trade) of 4-tuples (id, flavour, amt, time) where id is the *id* of the turtle that needs to be paid the associated flavour and amount at a given time (there is a global time for the model). This represents a set of time-indexed movements of apples and oranges to and from turtles, with a 'bank' holding payments until they are supposed to be delivered. Most trades lead to one, or two (if both parts of the trade are at non-zero delay), such entries going to the bank. Every time-step in the world the bank checks whether payments are due, and if so makes them to the agents in question, and removes the relevant entry from its list.

(4) Preliminary results

In order to test the model, we ran the *trading* component on its own, with no foraging, and each turtle being given an initial allocation of apples and oranges. Under that condition, the exponentially discounting turtles invariably end up holding almost all of the apples and oranges, and those of them that prefer apples have almost only apples, while those who prefer oranges have almost only oranges. How long this takes is dependent on the number of agents of each sort, and the values of some other parameters in the model. Nonetheless the process is

---

[5] It would have made no difference for the purposes of the present simulation to have allowed the turtles to go into 'debt'.

inexorable: a starting allocation with 200 hyperbolically discounting agents and only 2 exponentially discounting ones, where each individual holds the same intial stock, leads to a state of affairs where the two exponential discounters hold almost all (circa 98%) of the stock within 500 rounds of trading.[6]

A result *other* than one in which the exponential agents did better than non-exponential discounters in a trading only condition with the properties of our simulation would, of course, have to have been interpreted as an indication that our model was defective in some way. The hyperbolic discounters were *supposed* to be money pumped in the trading only condition, and the fact that they were indicated that the trading code we had written was behaving in one key respect as it should, and we could attempt to work out what happened when ecological factors came on line.

In the case of implementations of the model where foraging took place in the absence of trading it was less clear that a simple and powerful argument such as the money pump argument gave us a direct way of telling that the model was behaving as it should. Among the difficulties we faced here were:

(1) Not all of the hunches regarding the adaptiveness of hyperbolic discounting canvassed above have been tested, and those that have do not involve foraging problems fully isomorphic with our very simple one. That is to say, existing research was a poor guide to what to expect, so we were shooting in the dark.

(2) There are a significant number of dials that can be manipulated on even our very simple foraging model,[7] including:

- The discount rate for each kind of discounter.
- The number of hyperbolic discounters in the population.
- The number of exponential in the population.
- The number of ‘large’ patches in the model.
- The number of ‘small’ patches in the model.
- The size of the small patches.
- The size of the large patches.

In one sense the obvious way to use our model as a tool would have been to run it for various combinations of settings of the main parameters, so as to build up a phase portrait indicating under what regimes what sorts of discounting did better against opponents, and by what proportion. Doing that at all thoroughly for the set of parameters above would have involved a very large number of individual implementations of the model.

(3) The processing demands of the simulation for some ranges of parameter settings were significant. To give an indication, an early test run of a simpler model than the one described here in which only two parameters were varied (by about 50 small steps for each parameter) for runs of the simulation lasting less than 1000 time steps took around three full days of time on a high performance PC (each batch of settings was run 20 times). The ‘results’ themselves posed a further problem, consisting as they did in a file of comma separated values that contained several million numbers extracted from the set of runs of the model.

---

[6] At around 1500 rounds on a model with these settings the hyperbolically discounting agents typically held an average of 1 apple or orange each, and most were therefore practically excluded from any activity in the market since in our model fractional apples and oranges could not be traded.

[7] There are additional parameters that can be manipulated in the trading components of the model as well.

These three factors conspired to rule out, for the time being, anything like the comprehensive exploration that would otherwise be suggested. Further refinement of the model itself will come first, followed by some careful thinking about ways to streamling the process of exploring its dynamics.

With the above caveats noted, we can report the following. For classes of initial settings in which food patches significantly outnumber foraging agents, hyperbolically discounting agents tend to do better than exponentially discounting ones at our foragng problem, and this does not seem to depend on the relative discount rates of each type of agent. (More specifically, for a rough set of combinations of hyperbolic discounting rates as high as 5.0 and as low as 0.1, and for exponential discounting rates as low as 0.1 and as high as 0.95, the hyperbolic discounters always gathered 'food' from the patches at a higher rate.)

This 'result' was surprising to us. We'd expected that the different discounting dispositions built into the agents would make some difference, but for it to seem so clearly to favour one kind of discounter irrespective of their rate of discounting was unexpected. We therefore set out to construct an environment in which exponential discounters should do better (at a 'foraging' problem).

Reasoning that in a world where agents out-numbered or nearly out-numbered food patches, it might be possible to arrange the magnitudes of the large and small food patches and their relative density so that exponential discounters would always head for a very valuable large patch, whereas hyperbolic ones would occasionally get distracted by nearly small patches, and hence lose out in the longer run. A series of runs of the model with one of each kind of agent, one large patch (magnitude 100) one hundred small patches (magnitude 12) and highly skewed discount rates (0.01 for the exponential discounter, and 5.0 for the hyperbolic one)[8] had the predicted effect: exponential discounters can do systemmatically better even at a foraging problem, depending on the quantitative properties of the problem.

(5) Limitations and future work.

The work described here is, we emphasise again, very simple and very preliminary. Among the ways in which it could be extended are (a) in implementing standard 'stay or go' problems in the foraging space, and in varying patch quality as well as magnitudes, (b) in restricting the perception of the agents, so that there are gains from and risks to exploration, (c) in selecting a direction of travel according to functions besides making a beeline for the single most attractive option, but (say) towards the area with the greatest mean density of attractiveness.

A very preliminary conclusion from this initial exploration is that hyperbolic discounting may indeed be favoured in some foraging settings, but that no *general* argument for hyperbolic discounting should be trusted, since what sort of discounting is in fact favoured is dependent on properties of the enviroment.

In cases where there is an advantage to hyperbolic discounting, this may off-set the disadvantage due to vulnerability to money-pumping, as long as other factors constrain the rate at which agents can trade, or otherwise limit the rate of money pumping. In our own default model, the rate of stock transfer to exponential discounters is far faster than the gain to hyperbolic discounters when the world favours hyperbolic discounters at all. But in our default model a single time step sees one round of trading, and one spatial step in the foraging world. The ratio between these rates should itself be seen as a parameter in the model, and its significance explored.

---

[8] Subsequent exploration following confirmation of the intial hunch found that somewhat less extreme differences in discount rates preserved the effect.

References

Ainslie, G. (1992) *Picoeconomics*. Cambridge: Cambridge University Press.

Ainslie, G., & Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Animal Learning & Behavior*, 9, 476-482.

Ainslie, George and Haendel, V. (1983) The motives of the will In E. Gottheil, K. Druley, T. Skodola, H. Waxman (eds.), Etiology Aspects of Alcohol and Drug Abuse. Charles C. Thomas, pp. 119-140.

Axelrod, R. (forthcoming) Agent-Based Modeling as a Bridge Between Disciplines, in Kenneth L. Judd and Leigh Tesfatsion (Eds.), *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics,* Handbooks in Economics Series, North- Holland.

Forzano, L. B., & Logue, A. L. (1992) Predictors of adult humans' self-control and impulsiveness for food reinforcers. *Appetite,* 19, 33-47.

Grace, R. C. (1994) A contextual model of concurrent chains choice. *Journal of the Experimental Analysis of Behavior*, 61, 113-129.

Green, L., Fry, A., & Myerson, J. (1994) Discounting of delayed rewards: A life-span comparison. *Psychological Science*, 5, 33-36.

Kagel, J. H., Green, L., & Caraco, T. (1986). When foragers discount the future: Constraint or adaptation? *Animal Behaviour*, 36, 271-283.

Kirby, K. N. & Herrnstein, R. J. (1995) Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, 6, 83-89.

Kirby, K. N. & Marakovic, N. N. (1995) Modeling myopic decisions: Evidence for hyuperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes,* 64, 22-30.

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences*, 11, 665-709.

Mazur, J. E. (1997) Choice, delay, probability, and conditioned reinforcement. *Animal Learning and Behavior*, 25, 131-147.

Mazur, J.E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, *108*, 96-112.

Montague, P.R., and Berns, G. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36: 265-284.

Ross, D. (2005) *Microexplanation*. Cambridge, Mass. MIT Press.

Sobel, J.H. (2001) Money Pumps. *Philosophy of Science*, 68(2), 242-257.

Stephens, D. W. & McLinn, C. M. (2003) Choice and context: testing a simple short-term choice rule, *Animal Behaviour*, 66**,** 59–70.

Stephens, D. W., & Anderson, D. (2001). The adaptive value of preference for immediacy: when short-sighted rules have far-sighted consequences, *Behavioral Ecology* 12:330-339.

Stephens, D. W., Kerr, B. & Fernandez-Juricic, E. (2004) Impulsiveness without discounting: the ecological rationality hypothesis, *Proceedings of the Royal Society of London (B)* 271: 2459-2465.

Wilensky, U. (1999). NetLogo. (http://ccl.northwestern.edu/netlogo/). Center for Connected Learning and Computer−Based Modeling, Northwestern University, Evanston, IL.

Wilson, M. & Daly, M. (2003) Do pretty women inspire men to discount the future? *Proceedings of the Royal Society of London (Supplementary – biology letters),* 271, S177-179.

Stevens J.R., Hallinan E.V., & Hauser M.D. (2005) The ecology and evolution of patience in two New World primates. *Biology Letters,* 1:223-226.