Some Common Confusions about Hyperbolic Discounting

25 March 2008

Eric Rasmusen

Abstract

There is much confusion over what "hyperbolic discounting" means. I argue that what matters is the use of relativistic instead of objective time, not the shape of the discount function.

Dan R. and Catherine M. Dalton Professor, Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. visitor (07/08), Nuffield College, Oxford University. Office: 011-44-1865 554-163 or (01865) 554-163. Nuffield College, Room C3, New Road, Oxford, England, OX1 1NF. Erasmuse@indiana.edu. http://www.rasmusen.org. This paper: http://www.rasmusen.org/special/hyperbolic-rasmusen.pdf.

Keywords: Time inconsistency, hyperbolic discounting.

The term "hyperbolic discounting" comes up frequently nowadays in economists's lunchtime discussions.[1] Like "adverse selection" or "method of moments" it is one of those ideas that we all think we should know about. The idea of hyperbolic discounting, however, is easy to misunderstand. An example that I will explain below is that the term is usually applied to models that don't use hyperbolic functions for their discounting. There are other sources of confusion too, which I will try to clear up. This note will, I am sure, contain no single idea that is unknown to a great many scholars who have thought hard about hyperbolic discounting. I nonetheless hope that even those scholars could benefit from reading this, and could make use of this in explaining hyperbolic discounting to non-experts. For non-experts, of course, I hope this note clears away the most common misperceptions. First, I will explain what hyperbolic discounting is, and how it creates time inconsistency. Then, I will try to clear up various confusions about it.

What Hyperbolic Discounting Is

People prefer to consume more now rather than more later. The standard way to include this in economic analysis is using a positive personal discount rate in the utility function so consumption earlier adds more to utility than consumption later. An example of such a utility function is:

$$U_{2008} = C_{2008} + f(2009)C_{2009} + f(2010)C_{2010} + f(2011)C_{2011}, \qquad (1)$$

where the "discount function" is $f(t) < 1$ and f is declining in t. Or, we could write the discounting in terms of per-period "discount factors" $\delta_t < 1$, as in this example:

$$U_{2008} = C_{2008} + \delta_{2009}C_{2009} + \delta_{2009}\delta_{2010}C_{2010} + \delta_{2009}\delta_{2010}\delta_{2011}C_{2011}. \qquad (2)$$

We could but need not assume a constant discount factor δ and "discount rate" ρ, so that:

$$\delta_t = \delta^t = \left(\frac{1}{1+\rho}\right)^t. \qquad (3)$$

[1] Two overviews are the Frederick, George Loewenstein and Ted O'Donoghue 2002 article in the *Journal of Economic Literature* and the 2001 George-Marios Angeletos, David Laibson, Andrea Repetto, Jeremy Tobacman and Stephen Weinberg article in the *Journal of Economic Perspectives*. Those papers go into detail about motivations and applications for hyperbolic discounting, something I will not do here.

The functional form for discounting in equation (1) is perfectly general (though of course the linear utility separable across time is special). The functional form in equation (2), however, is an example of exponential discounting. Exponential discounting's key feature is that the time subscripts for the discount factors are absolute years, which I will denote by t, rather than relativistic "years in the future", which I will denote by τ (relativistic because they depend on the year in which the person starts). As some would put it, the difference is between the *date* t and the *delay* τ. Because of using relativistic discounting, if we view our person's decisions starting one year in the future, at 2009 instead of 2008, his utility function will be:

$$U_{2009} = C_{2009} + \delta_{2010}C_{2010} + \delta_{2010}\delta_{2011}C_{2011}. \tag{4}$$

The Time, t	0	1	2	3	4	5	6
The Discount Rate for $t-1$ to t, ρ	—	.10	.06	.05	.07	.07	.05
The Discount Factor for $t-1$ to t, δ_t	—	.909	.943	.952	.935	.935	.952
The Discount Function for 0 to t, $f(t)$	1	.909	.857	.816	.763	.713	.680

Table 1: Exponential Discounting (to 3 decimal places)

A nice property of exponential discounting is that a person's consumption path will be time consistent, meaning that if the person maximizes his utility at time 2000 by the choices $(C^*_{2008}, C^*_{2009}, C^*_{2010}, C^*_{2011})$ then he will

maximize them at time 2000 by the same values $(C^*_{2009}, C^*_{2010}, C^*_{2011})$ given his reduced wealth as the result of the consumption in 2008. The person may regret consuming so much in the year 2008, so his decisions across time are not consistent in the sense of being the choices he would make ex post, but the 2008 consumption is a sunk decision by 2009 anyway.

An alternative way a person might have time preference is to have discount factors whose levels depend not on the year itself– 2008, 2009, 2010– but on how many years in the future the consumption will occur–now, one year from now, two years from now. Looking at the decision made in year 0, the functional form could be exactly the same, e.g.

$$U_0 = C_0 + \delta_1 C_1 + \delta_1 \delta_2 C_2 + \delta_1 \delta_2 \delta_3 C_3, \tag{5}$$

where $\delta_\tau < 1$, using τ now instead of t because time is now relativistic rather than absolute. At time 1 (year 2009) however, the person would maximize $U_1 = C_1 + \delta_1 C_2 + \delta_1 \delta_2 C_3$, not $U'_1 = C_1 + \delta_2 C_2 + \delta_2 \delta_3 C_3$.

The function in equation (5) is just one of many ways time preference could be non-exponential, but whatever non-exponential form it takes, the person's decisions become time inconsistent. The optimal choices $(C^*_{2008}, C^*_{20091}, C^*_{2010}, C^*_{2011})$ from the 2000 utility function will not match the optimal choices using the 2001 utility function, $(C^{**}_{2009}, C^{**}_{2010}, C^{**}_{2011})$.

For example, it may be that the person is expecting a big income bonus in 2010. In year 2008, he might want to spread that income's consumption between 2010 and 2011 because though he highly values year 0 consumption, he is relatively indifferent between years 2 and 3. By the time 2010 arrives, however, year 2010 *is* year 0, and he would want to consume the entire bonus immediately.

Hyperbolic discounting is a useful idea for two reasons. First, it can explain revealed preferences that are inconsistent with exponential discounting. Second, and in my opinion more important, it can explain certain observed behaviors such as people's commitments to future actions when other explanations such as strategic positioning fail to apply, e. g., a person's joining a bank's saving plan which penalizes him for failing to persist in his saving.

What Hyperbolic Discounting Is Not

So much for what hyperbolic discounting is. What is it not? I will make the following points:

(a) Hyperbolic discounting is not about the discount rate changing over time. A constant discount rate is not essential for time consistency, nor does a varying discount rate create time inconsistency.

(b) "Hyperbolic discounting" does not, as commonly used, mean discounting using a hyperbolic function.

(c) Hyperbolic discounting really isn't about the shape of the discount function anyway.

(d) Hyperbolic discounting is not about someone being very impatient.

(e) Hyperbolic discounting is not necessarily about lack of self- control, or irrationality.

(f) Hyperbolic discounting does not depend delicately on the length of the time period.

(a) Hyperbolic discounting is not about the discount rate changing over time. A constant discount rate is not essential for time consistency, nor does a varying discount rate create time inconsistency.

Hyperbolic discounting makes not one but two changes from the standard model. First, it makes the per-period discount rate change over time. Second, it bases discounting on relativistic time rather than absolute time. It is this second assumption which is the key one. Both changes are necessary conditions to generate time inconsistency. I view the second change as the important one, though, because the assumption that the per-period discount rate is constant is merely a simplifying assumption, not meant to change the essential behavior of an economic model.

Comparison with interest rates on bonds is useful. We all know that long-term and short-term interest rates are different, and there is a large literature studying the term structure of interest rates, yet in most theoretical models we assume a constant rate of interest, r. We do that not because we think it realistic, but as a simplifying assumption, one that we often discard when it comes to actual estimation.

In the same way, there is no reason for us to think of a person's rate of time preference, ρ, as being constant, even in the standard exponential-discounting model with absolute time. If I am planning for the consumption of my 8-year-old daughter in 2008 I might use $\rho_t = 10\%$ for each year in the interval [2008, 2015] and then use $\rho_t = 5\%$ for the interval [2015, 2025] because I expect her degree of impatience to change. Doing so is still absolute-time discounting. It would become relativistic-time discounting only if in each of those 17 years I followed a policy of using 10% for whatever years were the 7 next years from the present and 5% for whatever years were the 8th to 17th year from the present.[2]

People change as they get older. It is not implausible that I would use a different rate of time preference at age 70 than at age 49. But that is fully compatible with exponential discounting. The key is that exponential discounting treats the parameter as "Rasmusen's rate of time preference for when he is 70 in the year 2058" whereas non-exponential discounting treats it as "Rasmusen's rate of time preference for 21 years from the present."

So long as the discounting uses absolute time, it will not give rise to time inconsistency, even if the discount rate changes over time. In 2008 I will be happy to have consumption from 2015 to 2016 discounted at a rate of 5%, and I will be equally happy with that plan in 2014.

It is also true that if my per-period discount rate is constant then it does not matter whether I discount using absolute time or relativistic time. If I use relativistic time, however, the assumption of a constant per-period discount rate is no longer just a simplifying assumption. It now becomes a substantive assumption, one necessary to avoid the changes in behavior that time inconsistency would bring.

(b) "Hyperbolic discounting" does not, as commonly used, mean discounting using a hyperbolic function. The standard way to model hyperbolic discounting is with the "quasi-hyperbolic discounting" of David Laibson (1997), which is simpler to use than a hyperbolic function and has similar

---

[2] We would not even expect impatience to always decline, in the standard model. Planning for myself, I might use $\rho_t = 3\%$ now, but $\rho_t = 15\%$ for the year 2058, when I expect to be in my second childhood and barely aware of future consumption.

properties, though it is not a special case of the hyperbolic function. Edward Phelps & Pollack (1968) had used the quasi-hyperbolic form to discount generations of people over time. Laibson (1997) applied it to within the self and named it "quasi-hyperbolic".[3] True "hyperbolic" discounting, proposed by Shin-Ho Chung & Robert Herrnstein (1961) in connection with a particular theory of behavior, has similar qualitative features but is more complicated to work with. Figure 1 shows how the shapes of the discounting over time differ.
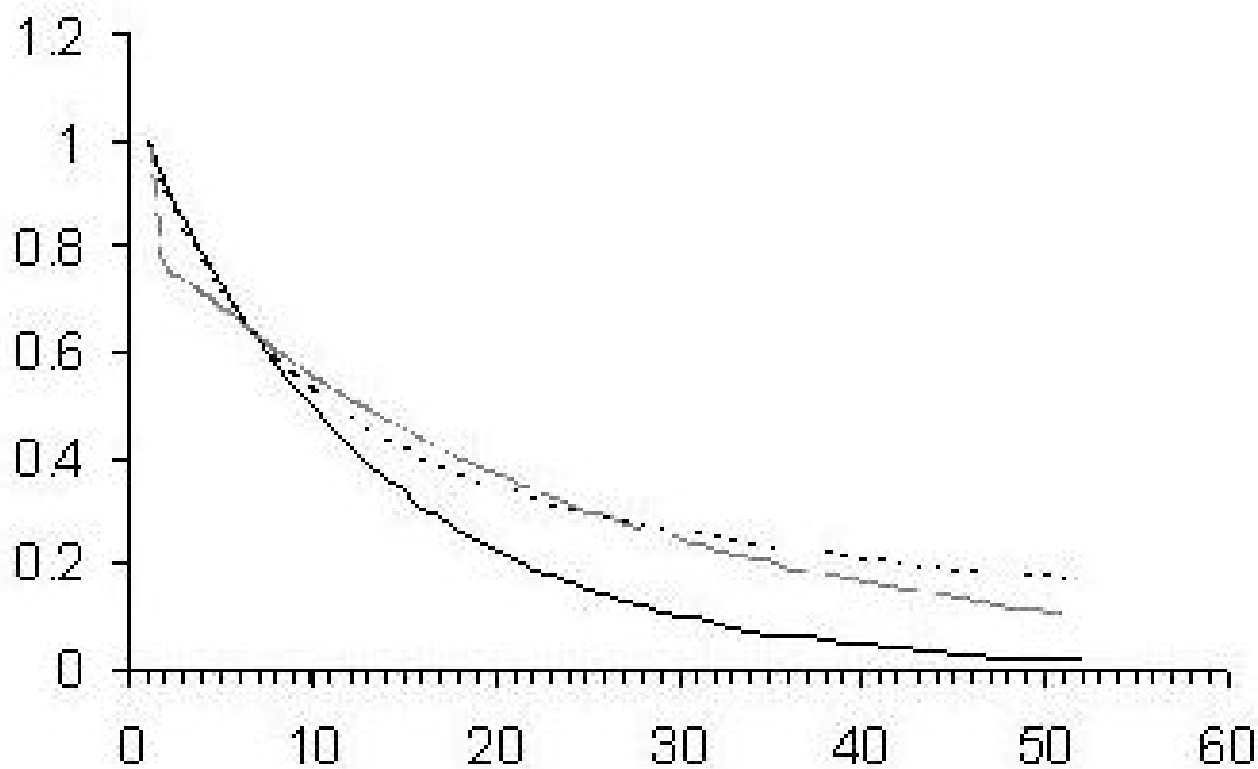

Figure 1:

---

[3]The term "subholic" has also been suggested, since the function lies below exponential and hyperbolic discounting for small values of τ– see Figure 1.

**The Shapes of Exponential (solid), Hyperbolic (dotted), and
Quasi-Hyperbolic (dashed) Discounting**
($\delta_{exp} = .92$, $f_h(\tau) = \frac{1}{1+.1\tau}$, $\beta = .8$ or $H = 1.25$ and $\delta_{qh} = .96$)

Exponential utility with a constant discount factor δ has the form $f(t) = \delta^t$:

$$U_0 = C_0 + \delta C_1 + \delta^2 C_2 + \delta^3 C_3 + ..., \tag{6}$$

which has the per-period discount rate $\rho = \frac{1-\delta}{\delta}$.

Quasi-hyperbolic utility (also called "Beta-Delta Utility") has the form $f(\tau) = \beta\delta^\tau$:

$$U_0 = C_0 + \beta\delta C_1 + \beta\delta^2 C_2 + \beta\delta^3 C_3 + ... \tag{7}$$

with $0 \leq \beta \leq 1$. This means that there is a lot of discounting from the present to time 1 (discount factor $\beta\delta < \delta$, and per-period discount rate $\rho = \frac{1-\beta\delta}{\beta\delta}$), but after that the discount factor between any two adjacent periods falls to a constant δ (which is the per-period discount rate $\rho = \frac{1-\delta}{\delta}$). And, of course, this is relativistic-time discounting, so the date of "the present" is always changing.

True hyperbolic utility has the form[4] $f(\tau) = \frac{1}{1+\alpha\tau}$:

$$U_0 = C_0 + \left(\frac{1}{1+\alpha}\right) C_1 + \left(\frac{1}{1+2\alpha}\right) C_2 + ... \tag{9}$$

Quasi-hyperbolic utility has some properties similar to hyperbolic utility, but it is not a special case of it.

As with quasi-hyperbolic discounting, what is most important is that hyperbolic discounting uses relativistic time in the utility function.

---

[4]Charles Harvey (1986) and George Loewenstein & Drazen Prelec (1992) proposed a more general form:

$$f(\tau) = \frac{1}{(1+\alpha\tau)^{\gamma/\alpha}}, \tag{8}$$

which has $f(\tau) = e^{-\gamma\tau}$ as its limiting case as α goes to zero.

All of these discounting functions are for use on the demand side, for the utility function. The supply side, the budget constraint, would be the same for all of them, and must use absolute time– t, not τ, in the notation here. Thus, even if the utility function is written in terms of τ, the budget constraint would be:

$$C_0 + \left(\frac{1}{1+r_t}\right)C_1 + \left(\frac{1}{1+r_1}\right)\left(\frac{1}{1+r_2}\right)C_2 + ... + \left(\frac{1}{1+r_1}\right)\left(\frac{1}{1+r_2}\right)\cdots\left(\frac{1}{1+r_t}\right)C_t \leq W_0, \tag{10}$$

where W_0 is the present value of wealth at time 0.

Any functional form for discounting except exponential discounting would give rise to time inconsistency. The particular functional form does not matter much, except for simplicity. This is important because one might think that hyperbolic discounting was a special case, which is false. Exponential discounting is the special case. It would be better to use the term "nonexponential discounting", but "hyperbolic" is what has become standard.

Hyperbolic discounting's functional form arises from the "Matching Law" theory of decisionmaking. Quasi-hyperbolic discounting amounts to adding to standard exponential discounting a parameter, β, that makes the person distinguish sharply between "Right now, time zero" and "All future times." Its functional form in equation (7) can also be written as:

$$U_0 = \gamma C_0 + \delta C_1 + \delta^2 C_2 + \delta^3 C_3 + ..., \tag{11}$$

with $\gamma > 0$, where $\gamma = 1$ for exponential utility, and where $\gamma > 1$ for quasi-hyperbolic utility. The marginal rate of substitution between consumption at time 0 and time τ' is γ/δ^τ, as opposed to the $(1/\beta)/\delta^\tau$ from equation (7), but they represent exactly the same consumer preferences.

I prefer form (11) to form (7), even though form (7) is standard. In form (11), the parameter γ represents what we might call "hyperbolicity" and allows γ to increase to infinity; whereas in form (7) the parameter β is declining in hyperbolicity, with a value restricted to be no greater than 1. Form (7) has the parameter β in every term, whereas in form (11) the parameter γ only has to show up once. For that reason, form (11) better represents the intuition that the parameter concerns the specialness of the present.

(c) Hyperbolic discounting really isn't about the shape of the discount function anyway. Rather it is about relativistic versus absolute time: how a person's discount rate for a given time period changes as that time period gets closer. At any one time, knowing the shape of the person's discount function for each period's future consumption doesn't tell you whether he is using "hyperbolic discounting" or not– except for the special case in which his per- period exponential discount rate is constant. This is really a reprise of point (b), but I will expand on it here using examples.

It is easy to come up with examples where exponential and non- exponential discount functions are identical.

Example. We will use the hyperbolic discounting function from Figure 1:

$$f(\tau) = \frac{1}{(1 + .1\tau)} \tag{12}$$

which generates the numbers in Table 2.

The Time: t or τ	0	1	2	3	4	5	6
The Discount Function for 0 to t or τ: $f(t)$ or $f(\tau)$	1	.91	.83	.77	.71	.67	.63
The Discount Rate for $t-1$ to t (or $\tau-1$ to τ): ρ_t or ρ_τ	—	10%	9%	8%	8%	7%	7%
The Discount Factor for $t-1$ to t, δ_t	—	.91	.92	.92	.93	.93	.94
The Quasi-Hyperbolic Delta Parameter for $t-1$ to t: δ_τ	—	.96	.92	.92	.93	.93	.94

Table 2: Exponential Discounting with a Hyperbolic Shape (to 2 decimal places, $f(\tau) = \frac{1}{(1+.1\tau)}$)

10

An exponential utility function with the same shape can be derived from $f(1) = \delta_1$, $f(2) = \delta_1\delta_2$, $f(3) = \delta_1\delta_2\delta_3$, ... $f(t) = \delta_1\delta_2\delta_3\cdots\delta_t$. so we can calculate

$$\delta_t = \frac{f(t)}{f(t-1)}, \tag{13}$$

and since $\delta_t = \frac{1}{1+\rho_t}$, we can calculate $\rho_t = \frac{1-\delta_t}{\delta_t}$.

Similarly, we can find a quasi-hyperbolic utility function with the same shape if we are allowed to vary the δ parameter. Let's set $\beta = .95$ We have $f(1) = \beta\delta_1$, $f(2) = \beta\delta_1\delta_2$, $f(3) = \beta\delta_1\delta_2\delta_3$, ... $f(\tau) = \beta\delta_1\delta_2\delta_3\cdots\delta_\tau$. so we can calculate

$$\delta_1 = \frac{f(1)}{\beta} \tag{14}$$

and

$$\delta_\tau = \frac{f(\tau)}{f(\tau-1)}. \tag{15}$$

In equation (15) and Table 2 all but the first of the δ parameters are the same for exponential and for quasi-hyperbolic discounting.

Note in Table 2 how the exponential discount *rates* are declining as time passes. This is a general feature of the hyperbolic discounting function with constant α. It is not a characteristic of the quasi-hyperbolic discounting function with constant δ, for which, of course, the discount factor is constant at δ after the first period so the discount rate is also constant.

(d) Hyperbolic discounting is not about someone being very impatient. A person can have high time preference even under standard exponential discounting. The key to hyperbolic discounting is that the person's high rate of discounting for a given future year's utility (say, utility in 2020) changes as that year approaches.

In theory, hyperbolic discounting could result in negative time preference, preferring future to present consumption. Someone might always care little about the present year, but a lot about future years. This would be one way to model a person who derives much of his utility from anticipation of future consumption. Patience of this kind would introduce time inconsistency too. In 2010 the person would want to consume a lot in 2015, but in

11

2015 he would prefer to defer consumption. Thus, the essence of hyperbolic discounting is not excessive impatience.

Figure 1 illustrates something else about the difference: although quasi-hyperbolic discounting discounts the close future more than exponential does, it discounts the distant future less. In Figure 1 the amount of discounting is about the same at year 6, but the exponential utility function is more patient at year 3 and less patient at year 30. Thus, the question of which kind of discounting is more patient is ill- formed.

(e) Hyperbolic discounting is not necessarily about lack of self-control, or irrationality. It is one way to model lack of self- control, to be sure, by having $0 < \beta < 1$ in the quasi-hyperbolic model. Lack of self-control is just an interpretation of the equations, though, not a necessary implication, and one might model lack of self-control in other ways. Also, non-exponential utility can be used to model other kinds of behavior not involving self-control. For example, we can imagine a person who is afraid of poverty more than 10 years in the future, but becomes braver as the poverty comes closer. Or, we could simply take the preferences at face value.

The question of whether in a particular setting hyperbolic discounting is being used to model (a) preferences that we usually don't assume in economics, or (b) mistakes such as lack of self-control, is important, especially for normative analysis. See Bernheim & Rangel (2008) or my own Rasmusen (2008) for two attempts to grapple with welfare analysis when discounting is hyperbolic. But although hyperbolic discounting does raise hard questions for welfare analysis, the hardest questions do not arise because it implies irrationality. Indeed, if we think of hyperbolic discounting as modelling irrationality, the welfare analysis gets easier– we simply try to undo the effect of the mistakes.

The question of whether hyperbolic discounting is rational is separate from the deeper question of whether discounting in general is rational or moral. This is an open question in moral philosophy. Most modern economists seem to agree that positive time preference is legitimate at the level of the individual; most moral philosophers seem to agree that it is illegitimate (and they even have doubts about discounting in general). See the references in Rasmusen (2008) for entry into the philosophy literature. The issue of what

the social discount rate should be is a separate one, which is even more contentious, in both disciplines.

(f) Hyperbolic discounting does not depend delicately on the length of the time period. Recall from equation (11) that a quasi-hyperbolic utility function can be written in the more intuitive form:

$$U_0 = H * C_0 + \delta C_1 + \delta^2 C_2 + \delta^3 C_3 + ...,$$

There is a big difference between the present and consumption at any future time, but the units in which time is measured do not affect tradeoffs between future time periods (though of course δ has to be written in the new time units too, so its value will change). For example, take a quasi-hyperbolic discount function with $\beta = .8$ and $\delta = .96$ and units of one year. The discount function will look like this:

Time:	0	1	2	3	4	5	6
Discount Function	1	.77	.74	.71	.68	.65	.63

Now change the time period to 2 years:

Time:	0	1	2	3	4	5	6
Discount Function	1	–	.74	–	.68	–	.63

The discount function values stay the same. We'd just change to $\delta_{new} = \delta_{old}^2 = .9216$ to match the new time period.

Concluding Remarks

Hyperbolic discounting is a useful idea for modelling people whose rate of time preference is based on relativistic time– distance in time from the present– rather than absolute time– the date. As with discounting using interest rates, we often model time preference with just one or two parameters but more realistically we should expect the rate of time preference to be constantly changing, whether we use exponential discounting or non-exponential. Such time preference is not necessarily irrational in the sense of being mistaken, but hyperbolic discounting can be used to model either someone whose preferences are based on relativistic time or someone who wishes he had more self- control. Either way, it gives rise to time inconsistency and can explain a variety of observed behaviors in which people try to commit, constraining their future selves.

References

Angeletos, George-Mariosm David Laibson, Andrea Repetto, Jeremy Tobacman & Stephen Weinberg (2001) " The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives,* 15(3): 47-68 (Summer 2001).

Bernheim, B. Douglas & Antonio Rangel (2007) "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics," NBER working paper 13737, http://www.nber.org/papers/w13737 (December 2007).

Chung, Shin-Ho & Richard J. Herrnstein (1961) "Relative and Absolute Strengths of Response as a Function of Frequency of Reinforcement," *Journal of the Experimental Analysis of Animal Behavior* 4: 267-272.

Frederick, Shane, George Loewenstein & Ted O'Donoghue (2002) "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2): 351-401 (June 2002).

Harvey, Charles M. (1986) "Value Functions for Infinite-Period Planning," *Management Science*, 32(9): 1123-1139 (September 1986).

Laibson, David (1997) "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2): 443-477 (May 1997).

Loewenstein, George & Drazen Prelec (1992) "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107(2): 573-597 (May 1992).

Phelps, Edward S. & R. A. Pollak (1968) "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies,*. 35: 185-199.

Rasmusen, Eric (2008) "Internalities and Paternalism: Applying Surplus Maximization to the Various Selves across Time," working paper, http://www.rasmusen.org/papers/internality-rasmusen.pdf (18 January 2008).

Strotz, R. H. (1956) "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23(3): 165-180 (1955-56).