BACKWARD DISCOUNTING

DEBRAJ RAY, NIKHIL VELLODI, AND RUQU WANG

ABSTRACT. We study a model in which lifetime individual utility is derived from both present and past consumption streams. Each of these streams is discounted, the former forward in the usual way, the latter backward. We presume that an individual at date t evaluates consumption programs according to some weighted average of her own felicity (as perceived at date t) and that of "future selves" at dates greater than t. This simple formulation allows agents to partially anticipate future regret in current decisions, and generates a set of novel testable implications in line with empirical evidence. The model is used to capture the notion of parental influence and investigate its impact on equilibrium savings. The paper also examines other applications of "backward discounting."

1. INTRODUCTION

In the standard model of intertemporal choice, individuals place weight solely on current and future outcomes. "Backward discounting" captures the notion that people also consider *past* outcomes in evaluating lifetime utility, discounted analogously to future outcomes. Under this postulate, lifetime individual happiness at any date is derived from current experience, anticipation and memory, with future and past discounted relative to the present. This paper discusses the implications of such a formulation.

There is an ostensibly simple reason why past utilities are typically left unconsidered.[1] The past is sunk and does not influence current decision-making. But there are several reasons to hold off on such an assertion.

Date: September 2017.

Ray: NYU and University of Warwick; Vellodi: NYU; Wang: Queen's University. Ray developed an early version of this idea under Guggenheim Fellowship, 1997-98. He acknowledges support under National Science Foundation Grant SES-1261560. Wang acknowledges financial support from SSHRC of Canada. We thank Gregor Smith, George Loewenstein, various seminar participants, and especially Matt Jackson, Ennio Stacchetti and Leeat Yariv.

[1]Strotz (1956) explicitly allows past outcomes to affect current utility. More recently, Gilboa, Postlewaite and Samuelson (2016) consider a model of memory utility in which a persistent effect of past consumption outcomes is present. Whilst our motivations have some similarities, we are primarily interested in the time-inconsistency that emerges from our formulation, whereas they study dynamically consistent consumption choices that might display less smoothing than in the standard model.

First, the perception of past experience may well influence current evaluation. For instance, suppose that you hire an agent to look after your business affairs over some years, and allocate a stream of consumable dividends to you. At the end of the relationship, you write the agent a letter of recommendation. Might your letter not depend on your utility evaluation of the agent's performance, as it appears to you at the *end* of this period, placing greater salience on recent performance? And if so, will that not affect the way in which the agent allocates consumption to you, knowing full well how recent outcomes influence you more strongly?

Second, when one agent has a say in decision-making for another, it might lead to conflict even though both agents fundamentally have the same preferences. Think of parenting: typically, a parent wants the child to be more forward-looking than the parent perceives the child as being. One possibility is that such disagreements emanate from distinct preferences, but this explanation begs the question of why those preferences are distinct to begin with.[2] Backward discounting hints at an alternative parsimonious explanation: a parent may want her son to be more forward-looking because she, in effect, is maximizing the utility of her child's *adult* self. Situated as she is later in life, she therefore places less weight on the child's early years and more on her later outcomes. Backward discounting then provides a natural mechanism for such considerations, without requiring differences in preferences from the outset.

Thirdly — and this is the central formulation studied in the paper — a person may place weight directly on future selves, who will look back on their lives.[3] Consider a 30 year old individual saving for retirement at age 70. What is she saving for? In part, she saves for the pleasures of retirement (consumption, status, charitable contributions, bequests). These are, of course, some distance away and are currently heavily discounted. But in part, the question is not *what* she is saving for but *whom* she is saving for: she is saving for her 70 year old self, and she may place a separate weight on the happiness of that self above and beyond the standard discount applied by her current self.

Now, the foregoing discussion may just appear to be semantic jugglery. Suppose that an individual has some "intrinsic" discount factor for future consumption, and additional concern for a "retirement self" or "shadow parent." Why not simply combine the two to arrive at some "effective" discount factor that evaluates future consumption? This

[2]See, e.g., Doepke and Zilibotti (2017) on models of parenting where parents and children are assumed to have different preferences.

[3]Ray (1987), Hori and Kanaya (1989), Pearce (2008) and Galperti and Strulovici (2017) also study models in which current generations place direct weight on the utilities of future selves, rather than on their consumption levels, without studying backward discounting.

is a perfectly sensible argument. The point is, however, that such discounting is *never* geometric, assuming that the "intrinsic" discount factors are geometric and that there is both forward and backward discounting. Indeed, we shall argue that our model gives rise naturally to some of the experimental observations that motivate models of hyperbolic discounting (see, e.g., Ainslie (1991), Loewenstein and Prelec (1992), Loewenstein and Thaler (1989), Laibson (1997, 1998), and O'Donoghue and Rabin (1999)). But our results are "minimal" in the sense we ask for no departures from the widely-used notion of geometric discounting. It is also in line with recent models of collective decision-making that give rise to time-inconsistency, such as Gollier and Weitzman (2010), Zuber (2011), and Jackson and Yariv (2015).

Moreover, our formulation generates additional testable predictions that are not made by the hyperbolic discounting model. As already discussed, individuals may prefer — at least ex post, and possibly ex ante to some degree — an improving sequence over a worsening sequence over time, where such a sequence might refer to diverse outcomes, such as relief from pain, emotional episodes, and monetary payments. These predictions are supported by numerous surveys and experiments.[4] Furthermore, our model generates a profile of local discount rates that are smaller in middle age, and greater both early on and in one's final years, a finding that finds empirical support; see, e.g., Harrison et al (2002) and Read and Read (2004).

That one parsimonious postulate can yield these disparate findings should be of some interest; see Proposition 1 for the details.

Why might a 30 year old also place weight on her 70 year old self? Why might a retirement self be salient in this way? One possible answer is that the 70 year old self is the imprint of *parental influence*, the result of many years of using a parent as a role model. When intergenerational links are strong (for instance in societies in which aged parents live with their grown children), there may be stronger empathy for the aged self. Section 3.4 draws out this interpretation in more detail. The theory we develop below suggests that societies with stronger parental influence or closer family ties will have higher savings rates and assign higher value to investment in human capital.[5] To

[4]See Loewenstein and Prelec (1993), Ariely (1998), Chapman (2000), Fredrickson and Kahneman (1993), Varey and Kahneman (1992), Loewenstein and Sicherman (1991), and the references in Ariely and Loewenstein (2000). We are grateful to George Loewenstein for providing us with the references.

[5]East Asia, where family ties are known to be relatively strong, is perhaps the region that would provide strong support for this point. Despite lower interest rates, the the savings rate in the region are higher than those for Western industrialized countries. According to World Bank (2015a), from 1980 to 2013, the average real interest rate (lending interest rate adjusted for inflation) is 3.7% for Japan, 3.8% for South Korea, and 1.9% for China, while it is 4.9% for the US, 2.8% for UK, 5.2% for France, and 8.1% for Germany.

develop these ideas further, we study the time-inconsistency embedded in the backward discounting model in the context of a standard life-cycle consumption problem. We show that sophisticated agents who correctly and strategically anticipate the consumption decisions of their future selves will engage in a higher rate of saving, one that exceeds the corresponding rate for the standard forward-discounting model. Indeed, greater weight placed on future selves results in higher savings rates, in line with the discussion above. We develop these arguments in Propositions 2–5. (All proofs are contained in the Appendix.)

Section 4 briefly discusses a number of diverse applications in which backward discounting might play a crucial role, such as addictive behaviour, political elections, and evolutionary selection, and raises some welfare considerations.[6]

In summary, what makes our model compelling is that a simple, plausible postulate of backward discounting, combined with placing weight on future selves, *simultaneously* generates a number of different findings and predictions in different economic and cultural contexts. We hope that the reader will view the current work as a first step to a deeper exploration of these matters.

2. Backward Discounting and Different Selves

2.1. **Preliminaries.** We set our model in continuous time. A person lives from period 0 to period N, and $\{c_s\}_{s=0}^N$ is a consumption path over the person's lifetime. Let u be an instantaneous utility function defined on consumption at every date.

Our main postulate is that a person at date t discounts both the past and the future. Consider a person currently alive at date t. She discounts future utilities at the local rate ρ_f, and past utilities at rate ρ_b. This forward and backward discounting system gives us the present value of the lifetime utility of her "current self" at date t as follows:

$$\int_0^t e^{-\rho_b(t-s)}u(c_s)ds + \int_t^N e^{-\rho_f(s-t)}u(c_s)ds$$

For the same period, the average gross national savings rate is 28.3% for Japan, 31.9% for South Korea, and 41.9% for China, while it is 16.3% for the US, 18.9% for UK, 21.9% for France, and 22.9% for Germany. (See also World Bank (2015b).) Hayashi (1997) tests various hypotheses explaining this significant difference in the savings rates using data on the US and Japan. One hypothesis that she mentioned but did not test is that the Japanese are more patient than the Americans. Our paper offers an explanation for why this may be true.

[6]As Caplin and Leahy (2004) observe, backward discounting has implications for the choice of the social discount rate. We come back briefly to this issue in the discussion, but do not pursue the welfare implications of backward discounting in this paper.

2.2. **Different Selves.** The discounting system in the above formula is only "piece-wise geometric" in s, and so generates time-inconsistency in the person's preferences. For example, a "current self" at date $t+a$ sees the present value of her lifetime utility as

$$\int_0^{t+a} e^{-\rho_b(t+a-s)}u(c_s)ds + \int_{t+a}^{N} e^{-\rho_f(s-t-a)}u(c_s)ds$$

which puts a weight of $e^{-\rho_b a}$ on $u(c_t)$ and 1 on $u(c_{t+a})$. In contrast, a "current self" at date t puts a weight of 1 on $u(c_t)$ and $e^{-\rho_f a}$ on $u(c_{t+a})$. Thus, the "current self" at date $t+a$ may regret the consumption choices made by the "current self" at date t.

Such regret may be partly anticipated by the person when she makes decisions. To capture this anticipation, the person places weight on the preferences of her "future selves." We describe this in more detail below, but briefly: we assume that an individual at date t places weight on two particularly salient selves: her "current self," as well as her "self" at some salient future date T. (She may place weight on other "intermediate selves" as well, but the added weight on a special future self is essential to our results.) The essential idea is that T represents a "rest-point" for the person — a special age at which she might feel most inclined to evaluate her life. It might be her retirement age, and we will call it that here, but it could be located elsewhere. For instance, it might be the imagined age of a "shadow parent," a parental role model who provides a natural position for the stock-taking of one's own life. It could well be N itself. It is also possible that T can change with current age t, and while we do not consider this case for expository convenience, it is entirely straightforward as long as we assume that $T-t$ is decreasing in t.

Observe that regret is quite endemic in this setting. The explicit consideration of future selves is crucial in allowing the agent to respond to such (anticipated) regret *by modifying her current choices*. Without placing weight on future selves, the regret caused by backward discounting would be a simple inevitability, not a matter to be accounted for in current decision-making.[7]

2.3. **The Profile of Patience.** We now proceed to a sharper formulation. Assume that at any date a person puts a weight of α on the preference of her current self, β on the preference of her retirement self T, and uniform weight ω on all other future selves. We make two observations on this specific formulation of the problem.

[7] To be clear, this interplay is not confined to single-agent problems. Recall the example of the populist government. In that setting, the government places weight on their future self at the re-election period, while voters discount backward.

First, while our individual places weight on all selves, clearly the current self and the retirement self have special roles, each receiving point mass. Indeed, we shall maintain the following assumption throughout the main text:

[A] $\rho_f \alpha$ and $\rho_b \beta$ both exceed ω.

Condition A endows the current and shadow-parent selves with privileged importance. The "flow equivalents" of their weights exceed the density on any of the other selves. In the special case with $\omega(t) = 0$, the model fits the shadow-parenting description explicitly. We use this simple form in the consumption-savings model of Section 3.

Second, we've assumed that the individual places weight on all future selves. Variations on this assumption are explored in the Online Appendix. For instance, the agent might place weight only on selves between t and T, or additionally on past selves as well as future selves. The outcomes are all qualitatively similar, as the variations in the Online Appendix reveal. At the same time, the presumption that there is a distinguished future self for whom (or around whom) the individual's weight spikes upward, *is* important. Without this presumption our main results are generally invalid; for more detail, see the Online Appendix.

Third, we've presumed that the weights are unchanged as the individual ages. One reason why this assumption could be dropped is that our individual might continue to place weights on selves who were "originally" in the future but "now" in the past, as she becomes older. The Online Appendix discusses this case as well,

With the above formulation in place, an individual located at date t will generate a payoff of $d(t, s)u(c_s)$ for consumption at date s, where $d(t, s)$ is an "effective discount factor" obtained by placing weight on discounts applied by all selves to date s. Lifetime utility for a person located at date t is therefore given by

$$\int_0^N d(t, s)u(c_s)ds. \tag{1}$$

We are particularly interested in these discount factors at pre-retirement ages $t < T$, for dates $s \geq t$. A little integration shows that these are given by:

$$d(t, s) = \alpha(t)e^{-\rho_f(s-t)} + \beta(t)e^{-\rho_b(T-s)} + \frac{\omega}{\rho_b}\left[1 - e^{-\rho_b(N-s)}\right] + \frac{\omega}{\rho_f}\left[1 - e^{-\rho_f(s-t)}\right]$$

when $t \leq s \leq T$, and by

$$d(t, s) = \alpha(t)e^{-\rho_f(s-t)} + \beta(t)e^{-\rho_f(s-T)} + \frac{\omega}{\rho_b}\left[1 - e^{-\rho_b(N-s)}\right] + \frac{\omega}{\rho_f}\left[1 - e^{-\rho_f(s-t)}\right]$$

when $T \leq s \leq N$. These discount factors can be translated to generate local rates of impatience — essentially, local discount rates — at any instant of time from the vantage point of any other instant. Specifically,

$$i(t,s) = \lim_{\epsilon \to 0} \Big[\frac{1}{\epsilon} \ln \frac{d(t,s)}{d(t,s+\epsilon)}\Big] = -\frac{d_s(t,s)}{d(t,s)}$$

measures the degree of "local impatience at s" that person t feels regarding choices at date s. In the standard model, $i(t,s) = \rho_f$.

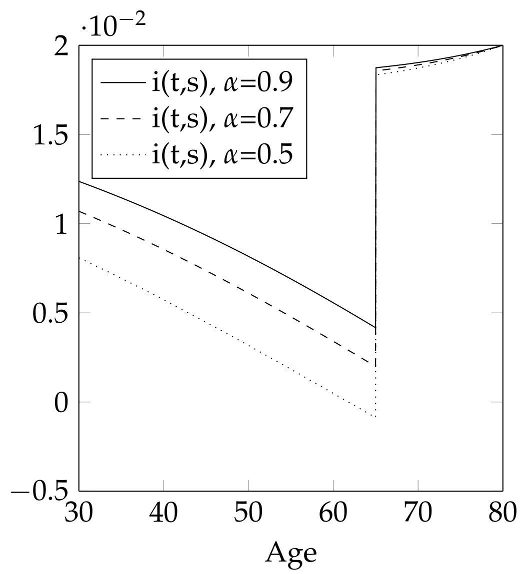
2.4. **Time Preference and Time Consistency.**

This simple model generates certain observed features in a natural way. First, unless the individual puts all weight only on the shadow parent at T, it creates time-inconsistency in individual decisions. This is almost immediate from examining the collection of effective discount factors $d(t,s)$. They are not geometric.

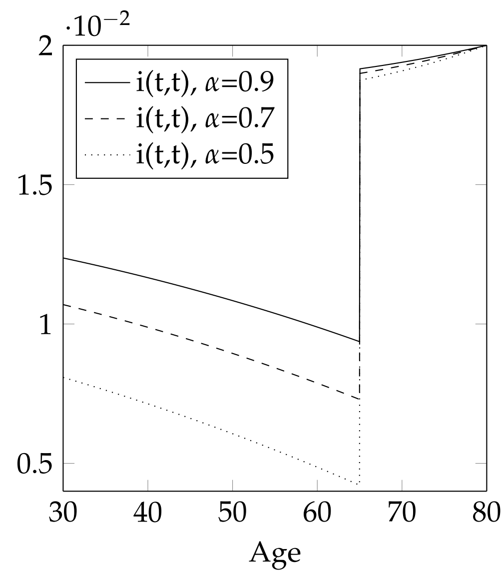
As an example, suppose that $\rho_b = \rho_f$, $t = 30$, $T = 70$, $\alpha(t) = \alpha$ and $\beta(t) = 1 - \alpha$, with no weight on any other selves. Notice that preferences over the very near future (say from age 30 to $30 + \epsilon$) are governed largely by the (exponential) discount factor of the *current* self — the comparison $1 : e^{-\rho_f \epsilon}$ matters far more than the comparison $e^{-40\rho_b} : e^{-(40-\epsilon)\rho_b}$, the latter being the weights applied by the 65 year old shadow parent. However, as the delay is pushed into the future — say to a consumption comparison across ages 50 and $50 + \epsilon$ — the bias towards the present exerted by the current self is increasingly compensated for by the bias towards the future exerted by the future self, and the two effects cancel, *from the vantage point of the thirty-year old*.

Notice that the time-inconsistency is of a particular kind. The example suggests that there is a bias towards *current* consumption (once the decision time comes around). This well-documented phenomenon is often captured by postulating non-geometric preferences (Strotz, 1956). The best-known of these is the hyperbolic class of discount factors (see Laibson (1997, 1998) and Harris and Laibson (2001)). The hyperbolic class, indeed, accounts for the following empirical regularity: discounting seems to be more active for time delays that are situated in the immediate future, whereas delays situated at a more distant date are viewed more neutrally (see Thaler (1981), Ainslie (1991), Loewenstein and Prelec (1992), Laibson (1997) and O'Donoghue and Rabin (1999)). Our model exhibits this phenomenon by combining forward and backward discounting, generating a sequence of nongeometric "effective" discount factors.

But the model makes more predictions that go beyond hyperbolic discounting, as summarized in the following result.



(A) $i(t,s)$, various s



(B) $i(t,t)$, various t

FIGURE 1. Local and Instantaneous Rates of Impatience for $t = 30$, $\rho_f = \rho_b = 0.02$, $\beta = 0.3$, $\omega = 0.001$ and Various Values of α.

PROPOSITION **1.** [1] *At each unretired age $t < T$, local impatience $i(t,s)$ declines in future dates s over all $s \in [t, T)$;*

[2] *Indeed, if a person is "young enough" (t is small), local impatience $i(t,s)$ at future dates s close to but smaller than T fall below 0 — that is, discounting could turn negative.*

[3] *On the other hand, for each unretired age t, local impatience $i(t,s)$ jumps up as s crosses the threshold T;*

[4] *An unretired individual displays more instantaneous patience as she grows older; i.e., $i(t,t)$ is decreasing in t — though always positive — for t smaller than T. But her instantaneous impatience jumps up as she crosses retirement: $i(t,t)$ discontinuously increases at $t = T$.*

[5] *Post retirement, an individual's impatience continues to grow with age. It approaches ρ_f as the individual approaches the end of her life, and in indeed roughly equal to ρ_f throughout, if ω is small enough.*

The first part of the proposition states that from the vantage point of an unretired person at date t, the relative impatience across adjacent dates in the future declines as the future is made more distant, being highest for choices between "today" and "tomorrow". This is where our model looks like hyperbolic discounting, with time-inconsistent present bias. Item 1 translates precisely into what Loewenstein and Prelec (1992) call the *common difference principle*: "if a person is indifferent between receiving $x > 0$ and $y > x$ at some later time, ... then she or she will strictly prefer the better outcome if both outcomes are postponed by a common amount [of time]".

But Parts 2–5 of the Proposition go beyond hyperbolic discounting. While a person at date t is generally impatient (in a perfectly standard way) over adjacent periods in the vicinity of the present, she might actually exhibit *negative* discounting for periods well into the future. This is the assertion in Part 2 of the Proposition. From the vantage point of the present, a young person plans to make real sacrifices in her middle age to provide for her post-retirement self, but is unprepared to make those sacrifices just yet.

However, Part 3 of the Proposition states that as this future contemplated date s crosses the age of the shadow parent T, then discounting again reverts to its old ways. After all, both t and T will discount adjacent dates at s (which exceeds both T and t) in exactly the same way, which induces a discontinuity as s crosses T. A young person not only plans great sacrifices in middle age, she plans to enjoy them just after crossing retirement.

Parts 4 and 5 of the proposition states that these youthful anticipations are — to some degree, but not fully — realized. These parts concern *current* levels of impatience $i(t,t)$ for both unretired and retired individuals. The proclivity of a person to be impatient over current and immediately adjacent future choices is indeed attenuated for ages approaching the retirement age T, though it is always positive and never negative as the person's younger incarnation might optimistically anticipate.

Yet T is truly a watershed age: as the individual slips over into retirement, her instantaneous rate of impatience jumps upwards. The last part of the proposition states that thereafter, her instantaneous rate of impatience continues to climb, as she "runs out" of future selves to care about. Indeed, if ω is close to zero, there are no future selves to care about at all, and her impatience will be roughly geometric and equal to the forward rate of discount ρ_f. Figure 1 summarizes these findings graphically, as well as highlighting the obvious role that the current weight α plays - a greater relative weight on the current self leads to greater impatience throughout life, as the attenuating force of future selves becomes less powerful.

We briefly return to the hyperbolic discounting model, this time to highlight the differences. In a discrete-time setting, a generalized hyperbolic discount function is given by

$$d(t,s) = [1 + \mu(s-t)]^{-\gamma/\mu}.$$

for positive constants γ and μ (see Loewenstein and Prelec (1992) for an axiomatic derivation.) This discount function gives us

$$i(t,s) = \frac{d(t,s)}{d(t,s+1)} = \left(1 + \frac{\mu}{1+\mu(s-t)}\right)^{-\gamma/\mu}.$$

In line with Proposition 1, one-period impatience in a generalized hyperbolic discounting model, $i(t,s)$, is decreasing in s.[8] In contrast, however, $i(t,s)$ does not jump up later, as it does in the case of backward discounting (when s crosses T). That is, a person is more patient when a trade-off occurs in a more distant future, as long as the trade-off does not occur after the date of her future self. Also in contrast with Proposition 1, $i(t,t)$ is constant under generalized hyperbolic discounting, while it is typically decreasing in the backward discounting model. An individual becomes more and more patient in a backward discounting model, while her current impatience remains the same in a hyperbolic discounting model.

This is a rich set of predictions from a single sparse model. We already know that present bias finds some confirmation in the literature. But the other predictions find support as well.

Harrison et al. (2002) investigate instantaneous discount factors $i(t,t)$ in a sample of Danish individuals across a rich set of demographic indicators. They observe that retirement "is associated with a discount rate over 12 percentage points higher than otherwise," but at the same time, there is a general tendency for discount rates to come down with age. These results speak directly to Part 4 of our proposition, and are consistent with the predictions there.[9] Table 1 shows selected results from this study.[10] There is a decline of discount rates into middle age, rising again in old age, and with retired individuals discounting at a higher rate than the overall average.

Read and Read (2004) derive similar results in a UK sample. Their results (see their Tables 3 and 4) show that discount rates broadly follow a U-shaped profile — the young and the old discount more heavily than the middle-aged. They also document an additional result of direct relevance here. They find that the young appear to fit a hyperbolic discounting model, at least over the short-to-medium run, whereas the old fit a

[8]Qualitative remarks regarding one-period and local impatience are directly comparable.

[9]Sozou and Seymour (2003) present a model of discounting under uncertain external hazard rates (mortality rates) and biological ageing (decreasing fertility) that generates such a profile. Our model, while consistent with the profile and the predictions of Sozou-Seymour, also generate the other predictions in Proposition 1, which Sozou and Seymour do not obtain — and in fairness, did not intend to.

[10]The numbers are taken from their fully stratified results, which control for possibly correlated covariates.

Demographic	Estimate	Standard Error
Young, < 29	**28.7**	0.96
Middle, $30 - 40$	**28.4**	0.87
Middle, $41 - 50$	**25.1**	1.07
Old, > 50	**30.0**	1.26
Retired	**38.7**	1.03
All	**28.2**	0.54

TABLE 1. Selected findings from Table 3, Harrison et al. (2002).

standard geometric discounting function. Were our agent to place no weight on future selves beyond T, or little weight on selves other than the current and retirement selves, our model would predict precisely this behavior: beyond retirement age, the conflict of interests between the different selves disappears, reducing the model to the standard geometric setting, as described in Part 5 of our proposition.

3. Equilibrium Plans

We now turn to the examination of actual consumption and savings decisions. We have already taken note of the inherent time-inconsistency present in our model of backward discounting. A person at date t, far removed from date T, will (by Proposition 1), *plan* on saving higher and higher fractions of her income as she gets into middle age and approaches retirement. The problem, of course, is that such plans will not in general be fully implemented by her future selves.

The purpose of this section is to examine both the "intended" behavior of "naive" agents who solve the usual intertemporal optimization problem paying no heed to issues of consistency, and the "equilibrium" behavior of "sophisticated" agents who take the choices of future selves into account. The idea in this second case is to treat each self as a separate player in a dynamic game, and solve for the subgame perfect equilibria of this game. Beginning with the work of Strotz (1956) and Phelps and Pollak (1968), these issues been studied in models with "changing tastes," especially models of hyperbolic discounting; see, e.g., Laibson (1994), Harris and Laibson (2001) and Bernheim, Ray and Yeltekin (2015).

For this section, we work with a simpler formulation in which a person places weight on just two selves: α on her current self at date t, and $\beta = 1 - \alpha$ on a "shadow parent" located at some fixed $T \leq N$. We presume that the local utility indicator is logarithmic, and that there is a common discount rate of ρ for both forward and backward discounting. All these simplifications are made for expositional ease: the analysis that follows can be conducted for all constant-elasticity utility functions, for weights placed on other selves, and for rates that vary across forward and backward discounting.

We also keep the background model of asset accumulation as simple as possible. There is a constant interest rate of r on both borrowing and lending. In addition, the agent receives an an exogenous income stream $\{y_s\}$ over the dates $[0, N]$. Let F_t denote financial assets and M_t the present discounted value of current and future income earnings at date t, i.e. $M_t = \int_t^N e^{-r(s-t)} y_s ds$. Then, if c_t is consumed at date t, we have $\dot{F}_t = rF_t + y_t - c_t$ and $\dot{M}_t = rM_t - y_t$. It is immediate that total asset wealth at date t, i.e. $A_t \equiv F_t + M_t$, must evolve according to

$$\dot{A}_t = \dot{F}_t + \dot{M}_t = r(F_t + M_t) - c_t = rA_t - c_t. \tag{2}$$

An agent located at date t seeks to choose $\{c_s\}$ over dates in the interval $[t, N]$ to maximize the expression in (1). Noting that the decisions up to date t have already been made, this is tantamount to maximizing

$$\alpha \int_t^N e^{-\rho(\tau - t)} \ln c_\tau d\tau + (1-\alpha) \int_t^N e^{-\rho|\tau - T|} \ln c_\tau d\tau \tag{3}$$

3.1. **Planned Consumption and Savings.** First, we ignore the time inconsistency problem and simply map out the profile of optimal consumption and savings from the vantage point of any date; say, date 0. That is, we maximize (3) evaluated at $t = 0$ by choosing a fully committed consumption plan over the interval $[0, N]$. The solution is summarized in the following result; see the Appendix for a detailed proof:

PROPOSITION **2.** *The optimal consumption profile viewed from date 0 is given by*

$$c_t(A) = \left[\frac{\alpha e^{-\rho t} + (1-\alpha) e^{-\rho|T-t|}}{\alpha e^{-\rho t} a_t + (1-\alpha) p_t} \right] A \equiv \lambda_t A \tag{4}$$

for $A \geq 0$, where a_t and p_t in this expression are given by:

$$a_t = \rho^{-1} \left[(\rho - 1) e^{-\rho(N-t)} + 1 \right] \tag{5}$$

and

$$p_t = \begin{cases} \rho^{-1} e^{-\rho(t-T)} [(\rho - 1) e^{-\rho(N-t)} + 1] & \text{for } t > T \\ \rho^{-1} \left\{ [(\rho - 1) e^{-\rho(N-T)} + 1] + [1 - e^{-\rho(T-t)}] \right\} & \text{for } t < T, \end{cases} \tag{6}$$

With the help of these closed forms (5) and (6), we can explore the planned consumption ratio λ_t. Notice first that there is an intrinsic tendency for the ratio to drift upwards simply by virtue of the finite horizon nature of the problem. For instance, in the last instant, all of permanent income will be consumed. We can "benchmark" this drift by setting $\alpha = 1$ in the problem above, whereupon the situation reduces to a perfectly standard life cycle problem. The consumption ratio sequence for this problem, which we denote by $\bar{\lambda}_t$, can be computed by putting $\alpha = 1$ in (4) above to see that

$$\bar{\lambda}_t = \frac{1}{a_t}, \tag{7}$$

and then invoking (5) to obtain a more explicit formula. This benchmark helps us get a handle on the extent to which our model departs from the standard formulation. Using (4) and (7), form the ratio

$$\theta_t \equiv \frac{\lambda_t}{\bar{\lambda}_t} = \frac{\alpha + (1-\alpha)e^{-\rho(|t-T|-t)}}{\alpha + (1-\alpha)(p_t/a_t e^{-\rho t})}, \tag{8}$$

and observe that a value of $\theta_t = 1$ implies that at date t there is no difference between the consumption ratios predicted by the two models. On the other hand, if $\theta_t < 1$ then the model of backward discounting predicts a higher savings rate, and this effect is directly related to the amount by which θ_t falls below unity.

Note that at any date $t \geq T$, and given the assumption that an individual places weight on just two selves, there is no difference between the discounting exhibited by the current self and the future self. Indeed, if we substitute the value of a_t (from (5) and the value of p_t (from (6) for $t \geq T$) in (8), we see that $\theta_t = 1$ for all $t \geq T$. For such time periods, then, there is no discrepancy (in consumption ratios) between our model and the usual formulation. Post-retirement, we are essentially back to the standard model.

The more interesting comparison is for dates that are less than T. The following observation is critical: for all $0 \leq t < T$,

$$e^{-\rho(T-t)} < \frac{p_t}{a_t}. \tag{9}$$

This is easy enough to establish by direct computation, using (5) and (6). Combining (9) with (8), we must conclude that consumption ratios — including the ratio at the current date — are lowered for all periods up to the age of the retirement self. This in itself is not surprising, as the agent places some weight on the retirement self at all dates, so that her effective rate of impatience is always lower than ρ.[11] What is of greater interest

[11]It should also be clear that if the agent increases the weight on his future self, then the savings rate increases at each date.

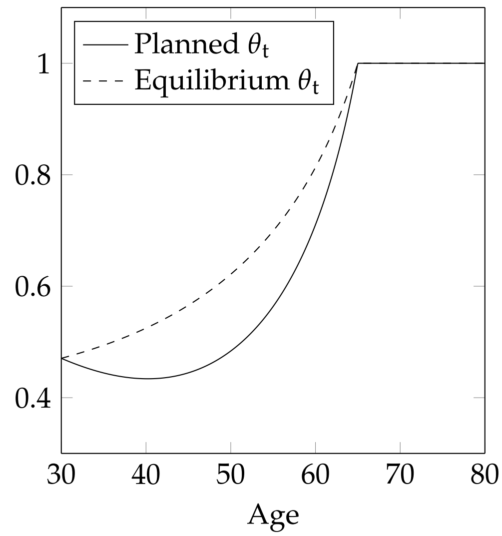


FIGURE 2. Planned and Equilibrium Savings Ratios Relative to the Standard Model, $\rho = 0.05$, $t = 30$, $N = 80$, $T = 65$, $r = 0.03$ and $\alpha = 0.5$.

is how the *extent* of that divergence changes with time. At what point over the agent's lifetime do we observe maximal (planned) divergence from the standard model?

It turns out, not surprisingly, that the answer to this question depends on the weight that the agent attaches to the future self. If this weight is high enough, the agent does most of her savings in the here and now, with consumption ratios rising over time relative to the standard model. Otherwise, the agent postpones the bulk of her savings to middle age, consuming more now and at retirement.

PROPOSITION **3.** *There exists $\hat{\alpha} \in (0, 1]$ such that if $\alpha \leq \hat{\alpha}$, θ_t always increases in t; while if $\alpha > \hat{\alpha}$, θ_t first decreases and then increases in t.*

Figure 2 illustrates the proposition with a numerical example. The solid line plots the discrepancy in the planned savings ratio relative to the standard model (θ_t) for $\rho = 0.5, r = 0.03$ and $\alpha = 0.5$. The agent's life starts at age 30, with a lifespan of 80 years and retirement planned for age 65. In this case, $\hat{\alpha} \approx 0.19$, so that any weight below 81% on the shadow parent leads to the U-shaped pattern in the discrepancy of savings ratios. That means that most savings — relative to the standard model — is planned for middle age.

3.2. **Equilibrium Consumption and Savings.** Now we take into account the possibility that planned paths will not be honored by future selves. As discussed previously, the idea is to treat each self as a separate player in a noncooperative game, and analyze the subgame perfect equilibria of such a game. The continuous time formulation, with a different player at every instant of time, is only a convenient limit formulation of the following natural discretization. Break the time horizon into discrete sub-intervals of length Δ. Allow each of these intervals to be controlled by only one agent. For that short interval, each agent is in effect solving a problem similar in nature to the planning problem described in Proposition 2. Consider an equilibrium profile thus generated and pass to the limit along a sequence of such equilibria as $\Delta \to 0$. We view the limiting profile as an equilibrium of the continuous time formulation. It turns out that there is a unique limiting profile:

PROPOSITION **4.** *The equilibrium consumption profile, i.e., the unique limit of equilibria in the discretization described above, is given by*

$$c_t^*(A) = \left[\frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha a_t + (1-\alpha)p_t}\right] A \equiv \lambda_t^* A \tag{10}$$

for $A \geq 0$, where a_t, p_t satisfy (5) *and* (6).

Notice that the solution to the a_t's and the p_t's are precisely what they were in the planned model. Any difference between the planned and equilibrium problems can be traced solely to a simple comparison of the formula in (10) with its predecessor (4). Recalling that consumption ratios in the standard model are given by $\bar{\lambda}_t = 1/a_t$, we can form our measure of "normalized" consumption ratios, just as we did before, by

$$\theta_t^* \equiv \frac{\lambda_t^*}{\bar{\lambda}_t} = \frac{\alpha + (1-\alpha)e^{-\rho|t-T|}}{\alpha + (1-\alpha)(p_t/a_t)}, \tag{11}$$

For dates that are less than T, (9) applies as before. Combining this inequality with (11), it is easy to see that consumption ratios —relative to the standard model — are again lowered for all dates up to the age of the retirement self. But we can say a bit more now than we could for the planning exercise: as long as $t < T$, θ_t *increases monotonically over time*. That is, the greatest impetus to the rate of savings — relative to the planning or full commitment model — comes at the earliest dates, no matter what the weight on the retirement self. To check this, it is sufficient to see that $e^{-\rho(T-t)}$ increases over time, while p_t/a_t declines, and to then apply these findings to (11).[12]

[12] It is immediate that $e^{-\rho(T-t)}$ increases with t. To check the claim for p_t/a_t, note that this expression equals $1 + \frac{1-x(t)}{(\rho-1)Dx(t)+1}$, where $x(t) = e^{-\rho(T-t)}$ and $D = e^{-\rho(N-T)}$. It is easy to check that this last expression is a monotonically decreasing function of x, while x is itself an increasing function of t.

It is worth reiterating that this equilibrium push towards early savings is not necessarily what an agent might want, if she could commit to a future plan. Indeed, Proposition 3 tells us that in general, this is *not* something that the agent wants. Rather, she would like future incarnations of herself to do the bulk of the savings. But given time-inconsistency, our individual knows that these large planned divergences are not going to be honored. That realization is built in into the equilibrium problem, generating the monotone trend that we've just described.

We summarize the foregoing discussion:

PROPOSITION **5.** *Suppose that an agent faces an accumulation problem, and places weights α on herself and $1-\alpha$ on a future self of fixed age T. Then*

[1] *In both the planning and equilibrium versions of the problem, consumption ratios (out of permanent income) are lower relative to those obtained for the standard problem, at every date $t < T$.*

[2] *For dates $t \geq T$, there is no discrepancy between the consumption ratios.*

[3] *A larger weight on the future self depresses consumption ratios even further (at each date), in both the planning and equilibrium versions.*

[4] *The planned consumption ratios are smaller (at each date t such that $0 < t < T$) than the equilibrium consumption ratios. Furthermore, the equilibrium consumption ratios monotonically increase over time, while no such presumption can be entertained for the planned divergence ratios (as described fully in Proposition 3).*

3.3. **A Numerical Example.** Just how large are the equilibrium effects in Proposition 5? Table 2 provides some numerical magnitudes for a particular parametric configuration. It describes equilibrium savings rates for individuals of age 30, 40 and 50, when lifetime is set to 80 and the age of the shadow parent set to a retirement age of 65. The interest rate on wealth is 3%, and the discount rate (in either direction) equals 0.01. Our model individual receives a flow income of 1 unit until retirement, and begins life at age 30 with an asset level of 2 units.[13]

To facilitate the use of everyday empirical observations, the savings rates we report are *not* out of permanent income, but out of current income (which is wage income plus any interest income on assets). The first column of the table reports savings rates at various ages in a standard life-cycle "benchmark model" (which is just our model

[13]Alternative specifications are available on request from the authors.

Age	Benchmark	Weight on Shadow Parent				
	Model	0.1	0.2	0.3	0.4	0.5
30	**34**	**37**	**39**	**42**	**44**	**47**
40	**27**	**29**	**31**	**33**	**35**	**37**
50	**16**	**17**	**19**	**20**	**21**	**23**

TABLE 2. Savings rates (%) out of current income in the benchmark and equilibrium models.

with α set equal to 1). The remaining columns add additional weight to the shadow parent. As expected, greater weight on the shadow parent results in a higher savings rate. Early on in life, this margin is particularly wide, with savings rates almost 50 percent higher under equal weights. The difference diminishes into old age, reflecting the convergence of consumption ratios as agents approach the retirement age. These numbers are provided, not necessarily for the sake of realism, but to suggest that socio-cultural phenomena such as upbringing and influence may have substantial effects on the proclivity to accumulate wealth.

3.4. **Some Remarks on the Shadow-Parent Interpretation of T.** It is well known that a parent's influence on their children is highly significant (see, e.g., Hess and Torney (1967), Bandera (1977) and Moschis (1987), among many others). Among other things, parents have a significant impact on such outcomes as their children's choice of career (Dryler (1998)), their focus on academic excellence (Salili (1994)), their perception of leadership (Harris and Hartman (1992)), their political attitudes (Hess and Torney (1967)), and on characteristics such as home ownership (Henretta (1984)). A smaller literature in economics (see, e.g., Becker and Mulligan (1997), Bisin and Verdier (1998, 2000) and Doepke and Zilibotti (2017)). For instance, Weinberg (2001) connects progeny behavior to the child-rearing practices of parents. Parents with higher incomes are more able to mold their children's behavior through pecuniary incentives, while parents with lower incomes have to rely on (less effective) nonpecuniary mechanisms, such as corporal punishment.

A particularly important target for parental influence is the attitude to consumption and savings. Moschis (1987, p. 77) summarizes the literature thus: "[T]here appears to be reasonably good supportive evidence that the family is instrumental in teaching young people basic rational aspects of consumption. It influences the development of rational consumption orientations related to a hierarchy of economic decisions delineated by previous writers. . . : spending and saving, expenditure allocation, and product decisions, including some evaluative criteria."

However, in our opinion, this form of influence extends far beyond *deliberate* attempts by parents to inculcate rudimentary notions of financial budgeting and other values in their children. The appropriate channel of influence may be more akin to what Hess and Torney (1967) have termed *anticipatory socialization*: the acquisition of attitudes and values about adult roles that have only limited relevance for the child but serves as a basis for subsequent adult behavior.[14] In an interesting and provocative essay, Brim (1966) views the socialization of an individual as a series of complex interpersonal relationships embedded in that individual. At the cost of some simplification, we might interpret this as stating that a particular personality is nothing more than the weighted combination of other personalities in the "cognitive neighborhood" of the individual in question. In a similar vein, the shadow parent in this model can also be replaced by "role models." People look up to their role models, who presumably are older in most cases. Furthermore, such role models often serve as implicit adjudicators of one's life choices, and as such, their effect on the current self might be well described by a mode of backward discounting. The weight of $1-\alpha$ on the "future self" then measures the influence of these role models.

Within the specifics of our formulation, the backward discounting model is a simple formulation in which a shadow parent situated at age 70, say, might disagree with the current self situated at age 30. We assume that all selves have the same preference *functional*, but that they discount both their future and their past. This gives rise to a minimal but cogent form of disagreement: the two selves will surely disagree on intertemporal decisions to be taken at intermediate ages. Parental influence refers to a state of affairs in which the shadow parent's preferences have been partly internalized by the current self.

4. Other Applications of Backward Discounting

There are several situations in which "backward valuations" might influence "forward decisions". This paper studied the implications of such an interaction on a standard life-cycle consumption problem. But there is a range of other situations where our formulation may be useful.

[14] As Ward (1974) describes it, such influence might be embodied in "implicit often unconscious learning for roles which will be assumed sometime in the future".

4.1. **Bad Habits and Policy Intervention.** We have seen how agents who place some weight on future selves might engage in more prudent behavior than standard forward discounters, foregoing current consumption to a greater degree. Turning from this positive angle to a normative one, a natural task might then be to identify contexts within which individuals could be *encouraged* to place greater weight on future selves, in order to improve lifetime outcomes. One such context is addictive behavior, for which the classic reference is Becker and Murphy (2004).[15] In their model, agents rationally trade off current consumption gains against future addiction losses. Specifically, given a stock S of past addictive consumption, local utility for consumption of normal goods c and addictive goods a is given by

$$U(c, a, S) = v(a, S) + u(c)$$

where $v_s(a, S) < 0$ captures the loss from addictive behaviour.[16] Standard policy interventions in this setting involve taxes (Becker, Grossman, Murphy (1994), Gruber, Köszegi (2001)). In such a context, the internalization of backward discounting via a concern for future selves would provide powerful incentives to attenuate addictive behavior, as future selves not only value present consumption less, but are manifestly harmed by the stock of addictive consumption. As such, an intervention that could serve to complement taxes might be to launch campaigns that serve to explicitly demonstrate the regret experienced by one's future self, thus combining both the future weight and backward discounting aspects of our model. One prominent example is that of 55 year-old Gerry Collins, diagnosed with terminal lung cancer, who appeared in numerous advertisements for Ireland's Health and Safety Executive, lamenting his past decisions. Anecdotal evidence estimate Collins' campaigns to have directly affected over 60,000 people to attempt to quit smoking.[17]

The situation is more complex when the activity in question is socially harmful but for which $v_s(a, S) > 0$. Examples might include miscreant behavior, such as crime or corruption. In this case, the presence of backward discounting might *exacerbate* the problem — agents now foresee the future rewards promised by such activities and thus consume at increased current levels.

4.2. **Social Decision-Making.** As Caplin and Leahy (2004) have observed, the phenomenon of backward discounting can give rise quite naturally to an argument for greater

[15] In their model of memory utility, Gilboa, Postlewaite and Samuelson (2016) discuss habit formation and addictive behavior within a rational, time-consistent framework.

[16] Note how, unlike the accumulation problem in Section 3, here the state variable enters utility *directly*.

[17] See the Health and Safety Executive of Ireland's 2014 press release.

patience on the part of a social planner, compared to the discount factor of the agents in the economy. Suppose that generation τ derives lifetime utility V_τ from a consumption stream $\{c_t\}$. A social planner might want to choose $\{c_t\}$ to maximize some weighted sum of these generational utilities,

$$\lambda_\tau V_\tau,$$

where the λ's are positive weights that integrate to unity. Assume that each generation exhibits a mixture of backward and forward discounting, so that:

$$V_\tau(\{c_t\}) = \int_\tau^\infty e^{-\rho_f(s-\tau)} u(c_s) ds + \int_0^\tau e^{-\rho_b(\tau-s)} u(c_s) ds,$$

where ρ_f is the forward rate and ρ_b is the backward rate. Then it is easy to see that the planner will exhibit a degree of patience that exceeds the patience implied in ρ_f. The planner will also be time-inconsistent, a problem in collective decision making that has been highlighted by Zuber (2011) and Jackson and Yariv (2015), among others.

4.3. **Evolutionary Models.** In many evolutionary situations, "success" or "fitness" is measured by some discounted sum of intertemporal rewards, where measurement takes place at the "end of the day". This might be particularly apt in the context of evolutionary biology, where fitness (over some intertemporal setting) is determined by the *end-state* value of some variable. Concentration on such end-state outcomes is equivalent to a very strong form of backward discounting in which the discount factor is set equal to zero.

Alternatively, in social or cultural contexts, "success" may be defined by the perceived lifetime rewards to an individual (or a role model, or a way of life) *at the end of the process*. If these perceptions form the basis for cultural selection, then agents (or behavior patterns) that assign heavier weight to future consumption would be the winners; in other words, agents (or cultural modes) that promote harder work and more savings when young are more likely to survive.

To be sure, there is no guarantee of such an outcome. It would all depend on which age groups (or which life stages) are considered natural points of success evaluation. These "natural" points are themselves subject to evolutionary or cultural pressures, leading to a deeper level of recursive analysis that is beyond the scope of this paper.

4.4. **Elections.** Suppose that a candidate can stay in office for at most two terms, and each term lasts for a time interval of length N. To simplify matters, assume that an elected officer's task is to provide a plan of allocating consumption over the time interval for which she is elected. Consider the first term. Let $V_f \equiv \int_0^N e^{-\rho_f(t)} u(c_t) dt$ denote the conventionally discounted total utility at time 0 of the median voter. A candidate's probability of being elected is then a function of V.

On the other hand, it is reasonable to suppose that the probability of being *re*-elected for a second term depends on the incumbent's perceived performance in the first term, and that voters may recall their past consumption using backward discounting. Let $V_b \equiv \int_0^N e^{-\rho_b(N-t)} u(c_t)dt$ denote the backward discounted total utility at the end of the first term. Suppose that a candidate can affect consumption paths. If she is forced to commit to consumption allocations in term 1 but cares about both terms, she would then be effectively maximizing some function $P(V_f, V_b)$, which is increasing in both of its arguments.

If the forward discount rate equals the interest rate on citizen assets, it is straightforward to verify that the constant consumption stream given by $c_t = A/(\int_0^N e^{-\rho_f(t)}dt)$ for all n maximizes V_f, where A is the present value of assets relevant to term 1.[18] That is, if a candidate does not aim for reelection, or if voters do not look back on their past experiences, a plan of constant consumption over time would be proposed.

Maximizing $P(V_f, V_b)$ subject to the same asset constraint implies the following equation: $P_1 e^{-\rho_f(n)} u'(c_t) + P_2 e^{-\rho_b(N-t)} u'(c_t) = \lambda e^{-rt}$, where $P_1 = \partial P(V_f, V_b)/\partial V_f > 0$, $P_2 = \partial P(V_f, V_b)/\partial V_b > 0$, and λ is an appropriate Lagrange multiplier. It follows that:

$$u'(c_t) = \frac{\lambda}{P_1 + P_2 e^{-\rho_b(N-t)} e^{\rho_f t}}$$

for all t (where we continue to assume that $\rho_f = r$). As the right hand side of the above equation is decreasing in t, and u is concave, we must conclude that $\{c_t\}$ is increasing over time. This reflects the observation that later consumption matters more to the voters at time of re-election.

If an elected candidate does not need to keep her promise, she would propose a plan of constant consumption to maximize V. But once elected, she would maximize V_b by reducing the consumption of early periods and increasing the consumption of later periods, so that her probability of being re-elected is maximized. The exact numbers can be calculated by setting $P_1 = 0$ in the above derivations. We can easily see that it makes the consumption profile even steeper, and thus reinforces the effects of re-election. Similar predictions would be generated by any model of career concerns, in which evaluations are carried out at the end of some initial term.

[18]There is an asset allocation problem over the longer run that we do not consider here.

5. Summary

In this paper, we introduce a framework in which lifetime individual felicities are derived from both present and past consumption streams. Each of these streams is discounted, the former forward in the usual fashion, the latter backward. We have argued that these two notions of discounting can interact in novel and interesting ways. In particular, we have shown that an individual who places weight on her current self as well as some future "retirement self" displays present bias in consumption. That is, she will exhibit impatience across alternatives that are positioned in periods adjacent to the present, but patience across similar choices positioned in the more distant future, so that lifetime choice plans are generally time-inconsistent much in the way predicted by the hyperbolic discounting model. But there are additional implications generated by the model: an individual gets more patient as she grows older during her working lifetime, but impatience jumps up again following the age of retirement. These results survive even if the individual cares for *all* future selves, as long as there is some sufficient weight on this particular stock-taking retirement self. We apply this model to consumption-savings decisions in a standard life-cycle model.

Similar considerations might arise for governments that make populist expenditures at the end of an election cycle. They know that the public might be "looking backward" to evaluate the government, discounting distant experiences more heavily while valuing recent outcomes. More generally, such considerations will arise whenever an agent (organization, individual) is evaluated by another: the latter may place greater weight on the agent's more recent activities. We have argued for similar effects in situations of cultural selection in which end states are viewed as ultimate measures of evolutionary success. The model also has potential implications for the study of parent-child interactions.

Our paper lies quite far from a general and comprehensive study of backward discounting. But we hope that it takes a small step in introducing the reader to the potentially important aspects of such a study.

6. Appendix

6.1. **Proof of Proposition 1. Part i**. Recall that the local rate of impatience at time s, from the vantage point of date t, is given by $i(t,s) = -d_s(t,s)/d(t,s)$. Simple differentiation of d shows that for all $s \in [t,T)$,

$$i(t,s) = \left[\frac{(\rho_f \alpha - \omega)e^{-\rho_f(s-t)} - \rho_b \beta e^{-\rho_b(T-s)} + \omega e^{-\rho_b(N-s)}}{(\alpha - \frac{\omega}{\rho_f})e^{-\rho_f(s-t)} + \beta e^{-\rho_b(T-s)} - \frac{\omega}{\rho_b} e^{-\rho_b(N-s)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}} \right] \tag{12}$$

By [A], we know that $\rho_b \beta - \omega > 0$, and hence $\rho_b \beta e^{-\rho_b(T)} - \omega e^{-\rho_b(N)} > 0$, and that $\rho_f \alpha - \omega > 0$. Hence $i(t,s)$ is a ratio of a decreasing function of s over an increasing one, and is therefore decreasing in s.

Part ii If the conditions of the statement hold, then $e^{-\rho_f(s-t)}$ can be made arbitrarily small, while $e^{-\rho_b(T-s)}$ approaches unity. Thus, the numerator of $i(t,s)$ becomes negative, while the denominator is always positive.

Part iii For $s \geq T$, note that

$$i(t,s) = \left[\frac{(\rho_f \alpha - \omega)e^{-\rho_f(s-t)} + \rho_f \beta e^{-\rho_f(s-T)} + \omega e^{-\rho_b(N-s)}}{(\alpha - \frac{\omega}{\rho_f})e^{-\rho_f(s-t)} + \beta e^{-\rho_f(s-T)} - \frac{\omega}{\rho_b} e^{-\rho_b(N-s)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}} \right]$$

To see that $i(t,s)$ jumps up at $s = T$, simply compare this expression to (12), and note that the only difference between these expressions at $s = T$ is the second term of the numerator, which turns from being strictly negative to strictly positive.[19]

Part iv For $t < T$, we can write $i(t,t)$ as

$$i(t,t) = \left[\frac{(\rho_f \alpha - \omega) - \rho_b \beta e^{-\rho_b(T-t)} + \omega e^{-\rho_b(N-t)}}{(\alpha - \frac{\omega}{\rho_f}) + \beta e^{-\rho_b(T-t)} - \frac{\omega}{\rho_b} e^{-\rho_b(N-t)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}} \right], \tag{13}$$

and once again, Condition A guarantees that the numerator is decreasing in t and the denominator is increasing. For $t > T$,

$$i(t,t) = \left[\frac{(\rho_f \alpha - \omega) + \rho_f \beta e^{-\rho_f(t-T)} + \omega e^{-\rho_b(N-t)}}{(\alpha - \frac{\omega}{\rho_f}) + \beta e^{-\rho_f(t-T)} - \frac{\omega}{\rho_b} e^{-\rho_b(N-t)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}} \right] \tag{14}$$

and again, we can easily compare (13) and (14) to confirm the positive jump in $i(t,t)$ as t crosses T.

[19] The two denominators are generally different as well, but at $s = T$ they are both the same.

Part v Differentiating (14) with respect to t, we see that $i(t,t)$ is increasing if $\rho_f(\rho_b + \omega) > \omega(\rho_b + \rho_f)e^{-\rho_b(N-t)}$, which holds by virtue of [A]. Finally, setting $\omega = 0$ yields the expression

$$i(t,t) = \left[\frac{\rho_f \alpha + \rho_f \beta e^{-\rho_f(t-T)}}{\alpha + \beta e^{-\rho_f(t-T)}}\right] = \rho_f$$

■

6.2. **Proof of Proposition 2.** An agent located at date 0 seeks to maximize, for any date t and any asset A_t at t,

$$e^{-\rho t}\alpha \int_t^N e^{-\rho s} \ln c_s ds + (1-\alpha) \int_t^N e^{-\rho|s-T|} \ln c_s ds \tag{15}$$

by choosing $\{c_s\}$, given some initial asset A_t and the law of motion (2) for assets. Let

$$V(A_t, t) = \int_t^N e^{-\rho(s-t)} \ln(c_s) ds \tag{16}$$

and

$$W(A_t, t) = \int_t^N e^{-\rho|T-s|} \ln(c_s) ds \tag{17}$$

be the values generated for each of the terms in the maximand by the supremum choice of consumption plan. That is, $V(A_t, t)$ records the forward discounted utility experienced from time t onwards, while $W(A_t, t)$ records the utility experienced at time t discounted backward and forward from time T, corresponding to the supremum choice of consumption plan.[20]

We guess and verify that these functions inherit the logarithmic utility structure, i.e. there exist coefficients a_t, b_t, p_t, q_t such that

$$V(A_t, t) = a_t \ln(A_t) + b_t \tag{18}$$

and

$$W(A_t, t) = p_t \ln(A_t) + q_t, \tag{19}$$

where the coefficients a_t and p_t satisfy the differential equations:

$$\begin{aligned} \rho a_t &= 1 + \dot{a}_t \\ \dot{p}_t &= -e^{-\rho|T-t|} \end{aligned} \tag{20}$$

with boundary values $a_N = 1$, $p_N = e^{-\rho(N-T)}$, and the coefficients b_t and q_t are continuously differentiable with respect to t, with $b_N = q_N = 0$.

[20] As the supremum value of the maximand is approached, it is easy to check that forward and backward utility integrals converge to well-defined limits, so that these values are well-defined.

With these guesses in hand, the time t problem (viewed from time 0) can be solved for all t. To do so, observe that the agent's supremum value from date t onward at initial asset A_t can be written as $e^{-\rho t}\alpha V(A_t,t)+(1-\alpha)W(A_t,t)$. Moreover,

$$e^{-\rho t}\alpha V(A_t,t)+(1-\alpha)W(A_t,t)=\sup_{\{c_s\}} e^{-\rho t}\alpha\Big[\int_t^{t+h} e^{-\rho(s-t)}\ln(c_s)ds+e^{-\rho h}V(A_{t+h},t+h)\Big]$$
$$+(1-\alpha)\Big[\int_t^{t+h} e^{-\rho|T-s|}\ln(c_s)ds+W(A_{t+h},t+h)\Big],$$

for all h such that $t+h\le N$, where the law of motion for A is given by $\dot{A}_s=rA_s-c_s$ for $s\ge t$. It follows that

$$\sup_{\{c_s\}} e^{-\rho t}\alpha\Big[\int_t^{t+h} e^{-\rho(s-t)}\ln(c_s)ds+e^{-\rho h}V(A_{t+h},t+h)-V(A_t,t)\Big]+$$
$$(1-\alpha)\Big[\int_t^{t+h} e^{-\rho|T-s|}\ln(c_s)ds+W(A_{t+h},t+h)-W(A_t,t)\Big]=0$$

for all such $h>0$. Dividing by h and taking the limit $h\to 0$, the above expression becomes

$$0=\sup_{c_t}\alpha e^{-\rho t}\big[\ln c_t+\dot{A}_t V_A(A_t,t)+V_t(A_t,t)-\rho V(A_t,t)\big]$$
$$+(1-\alpha)\big[e^{-\rho|T-t|}\ln c_t+\dot{A}_t W_A(A_t,t)+W_t(A_t,t)\big]$$

where $\dot{A}_t=rA_t-c_t$ and the subscripts on V denote suitable partial derivatives. (In taking this limit, we note that $V(A_t,t)$ and $W(A_t,t)$ are both differentiable, given their conjectured forms.[21]) Combining this expression with the formulae for $V(A_t,t)$ and $W(A_t,t)$ in (18) and (19), and taking first order conditions, we see that that the supremum is attained at

$$\alpha e^{-\rho t}\left[\frac{1}{c_t}-\frac{a_t}{A_t}\right]+(1-\alpha)\left[e^{-\rho|T-t|}\frac{1}{c_t}-\frac{p_t}{A_t}\right]=0,$$

so that

$$c_t=\left[\frac{\alpha e^{-\rho t}+(1-\alpha)e^{-\rho|T-t|}}{\alpha e^{-\rho t}a_t+(1-\alpha)p_t}\right]A_t=\lambda_t A_t \tag{21}$$

which verifies our asserted optimal policy at the instant t. But this is true only conditional on the conjectured forms of V and W, and the proof must now be completed by showing that V and W as well as their coefficients indeed have the forms (18), (19) and (20) for all dates, once we use the formula (21).

[21]Observe that a_t and b_t both have continuous derivatives. While p_t and q_t are possibly non-differentiable at T, (20) tells us that $\dot{p}_t$ is continuous at T, and hence that $\dot{q}_t$ is continuous at T.

With that in mind, define, for all t and all A_t

$$\hat{V}(A_t, t) \equiv \int_t^N e^{-\rho(s-t)} \ln(c_s(A_s))ds \text{ and } \hat{W}(A_t, t) \equiv \int_t^N e^{-\rho|T-s|} \ln(c_s(A_s))ds$$

where the family of functions $\{c_s(A_s)\}$ is given by (21). By totally differentiating with respect to t, it is easy to check that:

$$\rho\hat{V}(A_t, t) = \ln(c_t) + \dot{A}_t\hat{V}_A(A_t, t) + \hat{V}_t(A_t, t) \tag{22}$$

and

$$e^{-\rho|T-t|}\ln(c_t) + \dot{A}\hat{W}_A(A_t, t) + \hat{W}_t(A_t, t) = 0. \tag{23}$$

Optimality of the plan (21) will require us to equate (V, W) with $(\hat{V}, \hat{W})$, and that will verify both the conjectured functional forms in (18) and (19) as well as the differential equations (20) for the endogenous coefficients (a_t, b_t, p_t, q_t).

To derive the boundary values, set $t = N$ in (16) and (17), from which we see that $V(A_N, N) = \ln c_N$ and $W(A_N, N) = e^{-\rho(N-T)} \ln c_N$. To derive the ODEs for a_t and p_t, set $V = \hat{V}$ in (22) so that

$$\begin{aligned}
\rho V(A_t, t) &= \ln(c_t) + \dot{A}_t V_A(A_t, t) + V_t(A_t, t) \\
&= \ln(A_t) + \ln(\lambda_t) + \frac{\dot{A}_t a_t}{A_t} + \dot{a}_t \ln(A_t) + \dot{b}_t \\
&= (1 + \dot{a}_t)\ln(A_t) + \ln(\lambda_t) + \dot{b}_t + \frac{a_t(\dot{A}_t}{A_t}.
\end{aligned}$$

Using the law of motion $\dot{A}_t = rA_t - c_t$ in the equality above, we see after some trivial simplification that

$$\rho V(A_t, t) = (1 + \dot{a}_t)\ln(A_t) + \ln(\lambda_t) + \dot{b}_t + a_t(r - \lambda_t).$$

A similar calculation can be performed by setting $W = \hat{W}$ in (23) to obtain:

$$\begin{aligned}
0 &= e^{-\rho|T-t|}\ln(c_t) + \dot{A}_t W_A(A_t, t) + W_t(A_t, t) \\
&= e^{-\rho|T-t|}\ln(\lambda_t[A_t]) + \frac{p_t\dot{A}_t}{A_t} + \dot{p}_t\ln(A_t) + \dot{q}_t \\
&= (\dot{p}_t + e^{-\rho|T-t|})\ln(A_t) + e^{-\rho|T-t|}\ln(\lambda_t) + \dot{q}_t + \frac{p_t(\dot{A}_t)}{A_t} \\
&= (\dot{p}_t + e^{-\rho|T-t|})\ln(A_t) + e^{-\rho|T-t|}\ln(\lambda_t) + \dot{q}_t + p_t(r - \lambda_t),
\end{aligned}$$

Now use the formula for V in (18) and for W in (19) and equate coefficients to obtain the desired ODEs for $\{a_t, p_t\}$.

In the process we also obtain differential equations for $\{b_t, q_t\}$:

$$
\begin{aligned}
\rho b_t &= \ln(\lambda_t) + a_t(r - \lambda_t) + \dot{b}_t \\
\dot{q}_t &= p_t(\lambda_t - r) - e^{-\rho|T-t|}\ln(\lambda_t)
\end{aligned} \tag{24}
$$

from which it is readily verified that these coefficients satisfy the properties stated in the initial conjecture. All that remains is to provide explicit and unique solutions for a_t and p_t, which we can easily do given (20) and the boundary conditions. ■

6.3. **Proof of Proposition 3.** Let $D = e^{-\rho(N-T)}$, $x(t) = e^{-\rho(T-t)}$, $f(x) = x^2/e^{-\rho T}$, and

$$
g(x) = \frac{(\rho-1)e^{-\rho(N-t)} + 1 + 1 - e^{-\rho(T-t)}}{e^{-\rho t}[(\rho-1)e^{-\rho(N-t)} + 1]} = xe^{-\rho T} + \frac{1-x}{[(\rho-1)D + x^{-1}]e^{-\rho T}}
$$

Then

$$
\theta_t = \frac{\alpha + (1-\alpha)f(x(t))}{\alpha + (1-\alpha)g(x(t))} \equiv h(x(t)).
$$

It is straightforward to verify that $g''(x) < 0$. Furthermore, $h'(x) = 0$ implies that $f'(x)/g'(x) = h(x)$, which in turn implies that $g'(x) > 0$ whenever $h'(x) = 0$. Making use of all of the above, we see that $h''(x) > 0$ whenever $h'(x) = 0$. This indicates that $h(x)$ has at most one stationary point, and if that exists, $h(x)$ is decreasing in x to the left of that point and increasing to its right.

More tedious calculations show that at $x = e^{-\rho T}$ (i.e., at $t = 0$), we have $f(x) < g(x)$, while at $x = 1$ (i.e., at $t = T$), $f(x) = g(x)$. Therefore $h(x(0)) < h(x(T)) = 1$. Furthermore, $f'(x)g(x) - f(x)g'(x) > 0$ at $x = e^{-\rho T}$. Hence,

$$
\begin{aligned}
h'(x) = {} & \frac{1-\alpha}{[\alpha + (1-\alpha)g(x)]^2} \times \\
& \{f'(x)g(x) - f(x)g'(x) + \alpha[f'(x) - g'(x) - f'(x)g(x) + f(x)g'(x)]\}
\end{aligned}
$$

is positive at $\alpha = 0$, and possibly turns negative for larger α's if $f'(x) - g'(x) - f'(x)g(x) + f(x)g'(x) < 0$. The initial slope of $h(x)$ determines whether $h(x)$ is U-shaped or always increasing. If the initial slope is negative, then it is U-shaped. Otherwise, it is always increasing.

Since $x = x(t)$ is an increasing function of t, the same property holds for θ_t. ■

6.4. **Proof of Proposition 4.** The proof is completed in the following steps:

(i) Divide the full interval $[0, N]$ into sub-intervals of length Δ.

(ii) Solve the problem facing each agent at the initialization t of an interval, who controls consumption over $[t, t+\Delta)$, given some starting asset A, and under the restriction that total asset holdings at the end of the period must equal some provisionally exogenous value $\hat{A}$, feasible given A. The solution is an optimal plan as in Proposition 2.

(iii) Next, for each t and initial asset A, solve for the equilibrium choice of $\hat{A}$, under the conjecture that current and future agents behave as in (ii). This problem is in effect a game played by finitely many agents each taking a single action, and can be solved via backward induction.

(iv) Compute the limiting strategy profile as $\Delta \to 0$.

To this end, fix a t at the start of any sub-interval of length Δ, as well as "starting" and "ending" assets as described in (ii) above, and consider the problem faced by the agent controlling the interval $[t, t+\Delta]$. This problem is no different from our planning problem, except that (t, Δ) takes the place of $(0, N)$, and there is an asset promise of $\hat{A}$. But the latter only means that the effective present value of the asset at the agent's disposal is $B \equiv A - e^{-r\Delta}\hat{A} \geq 0$.[22]

It follows that for each of these "mini-problems," the solution is exactly as given by the planning problem: for any $t \in \{0, \Delta, 2\Delta, \ldots, N-\Delta\}$, and for $s \in [t, t+\Delta]$,

$$c_t(A_s, s; \hat{A}) = \left[\frac{\alpha e^{-\rho(s-t)} + (1-\alpha)e^{-\rho|T-s|}}{\alpha e^{-\rho(s-t)} a_s^t + (1-\alpha) p_s^t}\right] [A_s - e^{-r(t+\Delta-s)}\hat{A}] \tag{25}$$

where $\{a_s^t, p_s^t\}$ solve equations analogous to (5) and (6) for the planning problem:

$$a_s^t = \rho^{-1}\left[(\rho-1)e^{-\rho(t+\Delta-s)} + 1\right] \tag{26}$$

and

$$p_s^t = \begin{cases} \rho^{-1}\left[e^{-\rho(s-T)} + (\rho-1)e^{-\rho|(t+\Delta)-T|}\right] \text{ for} & s > T \\ \rho^{-1}\left[-e^{-\rho(T-s)} + (\rho+1)e^{-\rho|T-(t+\Delta)|}\right] & s < T, \end{cases} \tag{27}$$

These expressions provide a full characterization of the optimal consumption plan for the agent controlling $[t, t+\Delta)$, and aiming for $\hat{A}$ at the end of this period.

[22]Because $\hat{A}$ is feasible given A, B must be nonnegative.

There are also forward and backward values generated by the mini-plan, and these are exactly as in the planning problem, except that they are defined on $A - e^{-r\Delta}\hat{A}$:

$$V_t(A, \hat{A}) = a_t^t \ln(A - e^{-r\Delta}\hat{A}) + b^t, \tag{28}$$

and

$$W_t(A, \hat{A}) = p_t^t \ln(A - e^{-r\Delta}\hat{A}) + q^t, \tag{29}$$

where the additive constants b^t and q^t can be solved just as in the planning problem (see (24)), but we will not need to do so here.

Now we turn to the discrete game in which each of the players also choose the target $\hat{A}$ at the end of their control, anticipating a continuation payoff from that choice. We proceed by induction. At $N - \Delta$, define

$$V^*_{N-\Delta}(A) = \int_{N-\Delta}^{N} e^{-\rho(s-[N-\Delta])} \ln c_s ds$$

and

$$W^*_{N-\Delta}(A) = \int_{N-\Delta}^{N} e^{-\rho|T-s|} \ln c_s ds$$

to be the values generated by the planning problem at the very last decision stage, starting at A with a continuation asset of 0. It is easy to see that

$$V^*_{N-\Delta}(A) = \gamma_{N-\Delta} \ln(A) + \text{constant}$$

and

$$W^*_{N-\Delta}(A) = \nu_{N-\Delta} \ln(A) + \text{constant},$$

where

$$\gamma_{N-\Delta} = a_{N-\Delta}^{N-\Delta} = \frac{1}{\rho}[(\rho - 1)e^{-\Delta\rho} + 1]$$

and

$$\nu_{N-\Delta} = p_{N-\Delta}^{N-\Delta} = \frac{1}{\rho}[e^{-\rho|N-\Delta-T|} + (\rho - 1)e^{-\rho|N-T|}],$$

using (26) and (27). (The exact value of the constants need not concern us.)

Inductively, suppose that at decision date $t + \Delta$,

$$V^*_{t+\Delta}(A) = \gamma_{t+\Delta} \ln(A) + \text{constant} \tag{30}$$

and

$$W^*_{t+\Delta}(A) = \nu_{t+\Delta} \ln(A) + \text{constant}, \tag{31}$$

where $\gamma_{t+\Delta}$ and $\nu_{t+\Delta}$ satisfy the difference equations

$$\begin{aligned}\gamma_{t+\Delta} &= a_{t+\Delta}^{t+\Delta} + e^{-\rho\Delta}\gamma_{t+2\Delta} \\ \nu_{t+\Delta} &= p_{t+\Delta}^{t+\Delta} + \nu_{t+2\Delta}.\end{aligned} \tag{32}$$

The agent at date t chooses $\hat{A}$, for each initial asset A, to maximize

$$\alpha V_t(A, \hat{A}) + (1-\alpha)W_t(A, \hat{A}) + \alpha e^{-\rho\Delta} V_{t+\Delta}^*(\hat{A}) + (1-\alpha)W_{t+\Delta}^*(\hat{A}),$$

which, using (28), (29), (30) and (31), is equivalent to maximizing

$$[\alpha a_t^t + (1-\alpha)p_t^t]\ln[A - e^{-r\Delta}\hat{A}] + \alpha e^{-\rho\Delta}\gamma_{t+\Delta}\ln(A) + (1-\alpha)\nu_{t+\Delta}\ln(A)$$

by choosing $\hat{A}$. It is easy to see that the solution is given by

$$\sigma_t(A) = e^{r\Delta}\Big[\frac{\alpha e^{-\rho\Delta}\gamma_{t+\Delta} + (1-\alpha)\nu_{t+\Delta}}{\alpha(a_t^t + e^{-\rho\Delta}\gamma_{t+\Delta}) + (1-\alpha)(p_t^t + \nu_{t+\Delta})}\Big]A$$

where $\{a_t^t, p_t^t\}$ solve (26) and (27). From this expression, we can obtain solutions for $V_t^*(A), W_t^*(A)$ that solve (30)–(32) at date t. In particular, using the solutions

$$\begin{aligned}\gamma_t &= a_t^t + e^{-\rho\Delta}\gamma_{t+\Delta} \\ \nu_t &= p_t^t + \nu_{t+\Delta}\end{aligned} \tag{33}$$

we can write $\sigma_t(A) = e^{r\Delta}\mu_t A$, where

$$\mu_t = \left[\frac{\alpha(\gamma_t - a_t^t) + (1-\alpha)(\nu_t - p_t^t)}{\alpha(\gamma_t) + (1-\alpha)\nu_t}\right].$$

Now the inductive step is complete.

To finish the proof of the Proposition, we substitute the equilibrium rule $\sigma_t(A)$ for $\hat{A}$ in (25), which yields an instantaneous equilibrium consumption policy at date t:

$$c_t^*(A) = c_t(A, \sigma_t(A)) = \lambda_t[A - e^{-r\Delta}(e^{r\Delta}\mu_t A)] = \lambda_t(1-\mu_t)A,$$

where

$$\lambda_t \equiv \frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha a_t^t + (1-\alpha)p_t^t}.$$

Now we pass to the limit as $\Delta \to 0$. Equations (26) and (27) tell us that $a_t^t = 1 + \mathcal{O}(\Delta^2)$ and $p_t^t = e^{-\rho|T-t|} + \mathcal{O}(\Delta^2)$, so that by (33),

$$\begin{aligned}\gamma_t &= 1 + (1-\rho\Delta)\gamma_{t+\Delta} + \mathcal{O}(\Delta^2) \\ \nu_t &= e^{-\rho|T-t|} + \nu_{t+\Delta} + \mathcal{O}(\Delta^2).\end{aligned}$$

Thus, as $\Delta \to 0$, γ_t and ν_t satisfy the ODEs $\rho\dot{\gamma}_t = 1 + \gamma_t$ and $\dot{\nu}_t = -e^{-\rho|T-t|}$. Furthermore, it is easily verified that $\lim_{\Delta\to 0}\gamma_{N-\Delta} = 1$ and $\lim_{\Delta\to 0}\nu_{N-\Delta} = e^{-\rho|T-N|}$. These limiting boundary conditions, combined with the ODEs for γ_t and ν_t above, are precisely those that characterized the coefficients a_t and p_t in equations (5) and (6) of Proposition 2 and 4. Therefore $\lim_{\Delta\to 0}(\gamma_t, \nu_t) = (a_t, p_t)$, and so

$$\begin{aligned}
\lambda_t(1-\mu_t)A &= \left[\frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha a_t^t + (1-\alpha)p_t^t}\right]\left[\frac{\alpha a_t^t + (1-\alpha)p_t^t}{\alpha(\gamma_t) + (1-\alpha)\nu_t}\right]A \\
&\to \left[\frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha + (1-\alpha)e^{-\rho|T-t|}}\right]\left[\frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha a_t + (1-\alpha)p_t}\right]A \\
&= \left[\frac{\alpha + (1-\alpha)e^{-\rho|T-t|}}{\alpha a_t + (1-\alpha)p_t}\right]A
\end{aligned}$$

as $\Delta \to 0$, as required.

■

References

AINSLIE, G. (1991), "Derivation of 'Rational' Economic Behavior from Hyperbolic Discount Curves," *American Economic Review* **81**, 334–340.

ARIELY, D. (1998), "Combining Experiences over Time: The Effects of Duration, Intensity Changes, and On-line Measurements on Retrospective pain Evaluations," *Journal of Behavior Decision Making* **11**, 19–45.

ARIELY, D. AND G. LOEWENSTEIN (2000), "When Does Duration Matter in Judgment and Decision Making," *Journal of Experimental Psychology* **129**, 508–523 .

BANDERA, A. (1977), *Social Learning Theory*, Englewood Cliffs, NJ: Prentice-Hall.

BECKER, G. AND C. MULLIGAN (1997), "The Endogenous Determination of Time Preference," *Quarterly Journal of Economics* **112**, 729–758.

BECKER, G., M. GROSSMAN AND K. M. MURPHY, (1994), "An empirical analysis of cigarette addiction," *American Economic Review* **84**, 396-418.

BERNHEIM, D. AND D. RAY (1987), "Economic Growth with Integenerational Altruism," *Review of Economic Studies* **54**, 227–41.

BERNHEIM, D., RAY, D. AND S. YELTEKIN (2015), "Poverty and Self-Control," *Econometrica* **83**, 1877–1911.

BREGMAN, G. AND M. KILLEN (1999), "Adolescents' and Young Adults' Reasoning About Career Choice and the Role of Parental Influence," *Journal of Research on Adolescence* **9**, 253–75.

BISIN, A. AND T. VERDIER (1998), "On the Cultural Transmission of Preferences for Social Status," *Journal of Public Economics* **70**, 75–97.

BISIN, A. AND T. VERDIER (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics*, **115**, 955–988.

BRIM, O. (1966), "Socialization through the Life Cycle," in *Socialization after Childhood* (O. Brim and S. Wheeler (eds), New York: John Wiley and Sons.

CAPLIN, A. AND J. LEAHY (2004), "The Social Discount Rate," *Journal of Political Economy* **112**, 1257–1268.

CHAPMAN, G. (2000), "Preferences for Improving and Declining Sequences of Health Outcomes," *Journal of Behavioral Decision Making*, **13**, 203–218.

DOEPKE, M. AND F. ZILIBOTTI (2017), "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," mimeo., Northwestern University.

DRYLER, H. (1998), "Parental Role Models, Gender and Educational Choice," *British Journal of Sociology* **49**, 375–398.

FREDRICKSON, B.L., AND KAHNEMAN, D. (1993), "Duration Neglect in Retrospective Evaluations of Affective Episodes," *Journal of Personality and Social Psychology* **65**, 44–55.

GALPERTI, S. AND B. STRULOVICI (2017), "A Theory of Intergenerational Altruism," *Econometrica* **85**, 1175-1218.

GILBOA, I. A. POSTLEWAITE AND L. SAMUELSON (2016), "Memorable Consumption," *Journal of Economic Theory* **165**, 414-455.

GOLLIER, C., AND M. WEITZMAN (2010), "How Should the Distant Future Be Discounted When Discount Rates Are Uncertain?" *Economics Letters* **107**, 350–353.

GRUBER, J., AND B. KÖSZEGI (2001), "Is Addiction Rational: Theory and Evidence," *Quarterly Journal of Economics*, **116**, 1261-1303.

HARRIS, C. AND D. LAIBSON (2001), "Dynamic Choices of Hyperbolic Consumers," *Econometrica* **269**, 935–957.

HARRISON, G. W., MORTEN I. LAU AND MELONIE B. WILLIAMS (2002), "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review* **92**, 1606-1617.

HARTMAN, S. AND O. HARRIS (1992), "The Role of Parental Influence in Leadership," *Journal of Social Psychology* **132**, 153–167.

HAYASHI, F. (1997), *Understanding Saving: Evidence from the United States and Japan*, Cambridge, MA: The MIT Press.

HENRETTA, J. (1984), "Parental Status and Child's Home Ownership," *American Sociological Review* **49**, 131–140.

HESS, R. AND J. TORNEY (1967), *The Development of Political Attitudes in Children*, Chicago: Aldine Publishing Company.

HORI, H. AND S. KANAYA (1989), "Utility Functionals with Nonpaternalistic Intergenerational Altruism," *Journal of Economic Theory* **49**, 241–265.

JACKSON, M., AND L. YARIV (2015), "Collective Dynamic Choice: The Necessity of Time Inconsistency," *American Economic Journal: Microeconomics*.**7:4**, 150-178.

LAIBSON, D. (1994), ?Self-Control and Saving.? Mimeo., Department of Economics, Harvard University.

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

LAIBSON, D. (1998), "Life-Cycle Consumption and Hyperbolic Discount Functions," *European Economic Review* **42**, 861–871.

LOEWENSTEIN, G. AND D. PRELEC, (1992) "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics* **107**, 573–97.

LOEWENSTEIN, G. AND D. PRELEC, (1993) "Preferences for Sequences of Outcomes," *Psychological Review* **100**, 91–108.

LOEWENSTEIN, G. AND N. SICHERMAN, (1991) "Do Workers Prefer Increasing Wage Profiles," *Journal of Labor Economics* **9**, 67-84.

LOEWENSTEIN, G. AND R. THALER (1989), "Anomalies: Intertemporal Choice," *Journal of Economic Perspectives* **3**, 181–193.

MOSCHIS, G. (1987), *Consumer Socialization: A Life Cycle Perspective*, Lexington, MA: D.C. Heath and Company.

O'DONOGHUE, T. AND M. RABIN (1999), "Doing it Now or Later," *American Economic Review* **89**, 103–124.

PEARCE, D. (2008), "Nonpaternalistic Sympathy and the Inefficiency of Consistent Intertemporal Plans," in *Foundations in Microeconomic Theory* **89**, ed. by M. O. Jackson and A. McLennan. Berlin: Springer, 213 231.

PHELPS, E. AND R. POLLAK (1968), "On Second-Best National Saving and Game Equilibrium Growth," *Review of Economic Studies* **35**, 185–199.

RAY, D. (1987), "Nonpaternalistic Intergenerational Altruism," *Journal of Economic Theory* **41**, 112–132.

READ, D. AND N.L. READ (2004), "Time Discounting Over the Lifespan," *Organization Behaviour and Human Decision Processes* **94**, 22-32.

SALILI, F. (1994), "Age, Sex, and Cultural Differences in the Meaning and Dimensions of Achievement," *Personality and Social Psychology Bulletin* **20**, 635–648.

SOZOU, P.D. AND R.M. SEYMOUR (2003), "Augmented Discounting: Interaction between aging and time-preference," *Proceedings of the Royal Society of London B* **270**, 1047-1053.

STROTZ, R. (1956), "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies* **23**, 165–180.

VAREY, C.A., AND D. KAHNEMAN (1992), "Experiences Extended across Time: Evaluation of Moments and Episodes," *Journal of Behavioral Decision Making* **5**, 169–185.

WARD, S. (1974), "Consumer Socialization," *Journal of Consumer Research* **1** (September), 1–16.

WEINBERG, B. (2001), "An Incentive Model of the Effect of Parental Income on Children," *Journal of Political Economy* **109**(2), 226–280.

WORLD BANK (2015a), "Real Interest Rate," *World Development Indicators, World DataBank*.

WORLD BANK (2015b), "Adjusted Savings: Gross Savings (% of GNI)," *World Development Indicators, World DataBank*.

ZUBER, S. (2011), "The Aggregation of Preferences: Can We Ignore the Past?" *Theory and Decision* **70**, 367–384.