Current draft: May, 1994.
First draft: April, 1992.
Comments welcome.

# Self-Control and Saving

David Laibson[1]

Research on animal and human behavior has led psychologists to conclude that discount functions are approximately hyperbolic (Ainslie, 1992). This paper characterizes the savings/consumption behavior of sophisticated economic decision-makers with hyperbolic discount functions. Such preferences are characterized by "dynamic inconsistency"; the preferences imply a conflict between the optimal contingent plan from today's perspective and the optimal decision from tomorrow's perspective. To model intertemporal choice when such conflicts arise, I assume that individuals engage in an intertemporal game with themselves (Strotz (1956), Phelps (1968), Peleg and Yaari (1973), Goldman (1980)). Specifically, I reinterpret the infinite-horizon intergenerational consumption game proposed by Phelps and Pollak (1968) as an intra-personal consumption game. I show that hyperbolic discounting generates a coordination problem which leads to the existence of multiple intra-personal equilibria, some of which are Pareto-rankable. (In this intra-personal game two equilibria are Pareto-rankable if all temporal selves of the individual are better off under one of the two equilibria.) I characterize the equilibrium set, and calibrate the model. I interpret this model as a framework for understanding the psychological concept of self-control.

[1]This work has been supported financially by the National Science Foundation and the Alfred P. Sloan Foundation. I have benefitted from the insights of Roland Benabou, Olivier Blanchard, Drew Fudenberg, Fausto Panunzi, Drazen Prelec, Matthew Rabin, Julio Rotemberg, Jean Tirole, Jeromin Zettlemeyer, and participants in the MIT Theory Lunch and the MIT Macroeconomics Workshop. All mistakes should be blamed on my period $t-1$ self.

# 1 Introduction

Research on animal and human behavior has led psychologists to conclude (see Ainslie (1992)) that discount functions are approximately hyperbolic: *e.g.*, rewards $\tau$ periods in the future are discounted $\frac{1}{\kappa_1+\kappa_2\cdot\tau}$ where the $\kappa$'s are constants. Hyperbolic discount functions imply a monotonically falling discount rate. This discount structure sets up a conflict between today's preferences and the preferences which will be held in the future. For example, from today's perspective, the discount rate between two far off periods, $t$ and $t+1$, is a long-term low discount rate. However, from the time $t$ perspective, the discount rate between $t$ and $t+1$ is a short-term high discount rate. This type of preference "change" is reflected in many common experiences. For example, today I may desire to quit smoking next year, but when next year actually roles around my taste at that time will be to postpone any sacrifices another year.

Hyperbolic discount functions generate a preference structure which is a special case of the general class of "dynamically inconsistent" preferences: *i.e.*, preferences which imply a conflict between the optimal contingent plan from today's perspective and the optimal decision from tomorrow's perspective. Robert Strotz (1956) was the first economist to study dynamically inconsistent preferences. Pollak (1968), Peleg and Yaari (1973), and Goldman (1980) have extended Strotz's work, arguing that when preferences are dynamically inconsistent, dynamic decisions should be modelled as an *intra*-personal game among different temporal selves (*i.e.*, "me today" is modelled as a different player from "me tomorrow").

Despite the availability of this analytic framework, and the substantial body of evidence supporting hyperbolic discounting, few economists have studied the implications of hyperbolic discount functions. Phelps and Pollak (1968) analyze an inter-generational game in which each generation has a discount function which I will argue below is approximately hyperbolic. However, Phelps and Pollak restrict their attention to linear Markov-equilibria, undermining the generality of their analysis. Akerlof (1991) analyzes the behavior of decision-makers who place a special premium on effort made in the current period. Such a premium can be interpreted as a reflection of hyperbolic discounting although this interpretation is not made by Akerlof. Akerlof's analysis is inconsistent with the intra-personal game approach as his decision-makers act myopically; they fail to foresee the preference "changes" described above. Finally, Loewenstein and Prelec (1992) also analyze the choices of myopic decision-makers with hyperbolic discount functions.

To date economists have not characterized the unrestricted behavior of sophisticated/rational economic decision-makers with hyperbolic discount functions. However, in pathbreaking work, a pschyiatrist, George Ainslie (1992), has discussed the kinds of qualitative behavior that sophisticated agents with hyperbolic discount functions might exhibit. The primary goal of this paper is to formalize and extend Ainslie's psychological analysis, by explicitly modelling individual decision-making as an intra-personal game.

Specifically, I adopt preferences which nest the standard exponential discount structure as a special case, and I analyze an intra-personal delay-of-gratification game which is associated with these preferences. I look for subgame-perfect equilibria of this game and find a multiplicity which is disconcerting because it arises *even* in the special case of exponential discounting. To address this problem, I propose a menu of three refinements, *each* of which eliminates the multiplicity when discounting is exponential. For the exponential case, moreover, each of these refinements admits a single equilibrium which is the standard "Ramsey" outcome.

For the case of hyperbolic discounting, only two of the original three refinements still imply uniqueness. A common equilibrium survives both of these refinements, but this equilibrium is Pareto-dominated by a continuum of other subgame perfect equilibria. (In this intra-personal game two equilibria are pareto-rankable if all temporal selves of the individual are better off under one of the two equilibria.) This suggests that there is no focal equilibrium in the non-exponential discount case. So while there are reasonable ways to rule out multiplicity in the exponential case, such arguments do not carry over to the case in which discount rates are non-exponential.

This conclusion has interesting consequences for the theory of dynamic choice in general and for consumption theory in particular. In the dynamic choice category there are three conclusions which should be highlighted. First, this approach provides a way of explaining why people with identical preferences might exhibit dramatically different behavior in the same environment. Second, this approach can explain why some people have self-acknowledged "bad habits" which they can't break, while other people with ostensibly identical preferences muster the internal self-discpline/self-control to avoid these "bad habits." The bad habit is a perfect equilibrium which is Pareto-dominated by another perfect equilibrium. Third, the model explains why individual behavior is often perceived to be self-diagnostic of future behavior (*i.e.*, the model explains why individuals often reason, "if I show self-control today, I'll be more likely to show self-control tomorrow").

Finally the model is of direct interest to macroeconomists, since the intertemporal choice problem that I study is a savings/consumption decision. The model explains how an economy's savings rate can be indeterminate in equilibrium. A single calibration of the model produces many potential equilibrium paths. For any reasonable parameter vector the associated range of equilibrium savings rates is quite large, with substantially more dispersion than that exhibited in cross-country savings studies. Hence, the multiple equilibria of this model can explain heterogeneous international savings rates. This suggests that the heterogeneity implied by the model is substantial when normalized by cross-country savings variability.

The paper is divided into six sections, including this introduction. Section two of the paper lays out the formal model, describing the specific intra-personal game that I consider. Section three characterizes the equilibrium set for the case of exponential discounting, and proposes three refinements. Section four applies the refinements of section three to the case of non-exponential discounting. Section five characterizes the equilibrium set when the discount structure is non-exponential. Section six interprets the results and section seven concludes.

# 2 An intra-personal game.

Following Strotz (1955), Pollak (1968), Peleg and Yaari (1972), and Goldman (1980), I model an individual as a composite of autonomous temporal selves. These selves are indexed by their respective periods of control, $(t = 0, 1, 2, \ldots)$, over a consumption decision. During its period of control, self $t$ observes all past consumption levels $(c_0, c_1, c_2, \ldots, c_{t-1})$, and the current wealth level $A_t$. Self $t$ then chooses a consumption level for period $t$, which satisfies the restriction,

$$0 \leq c_t \leq A_t. \tag{1}$$

Self $t+1$ then "inherits" an asset stock equal to,

$$A_{t+1} = R \cdot (A_t - c_t) \tag{2}$$

where $R$ is the constant gross return. The game continues, with self $t+1$ in control. Finally, note that in this game the precommitment solution discussed by Strotz (1956) is implicitly ruled out; *i.e.*, the current self can not make consumption decisions for future selves.[2]

Now it only remains to specify the payoffs of the "players" of this game. Player $t$ receives payoff $U_t(c_0, c_1, \ldots, c_t, \ldots)$ where $U_t$ is a map into the real line or the extended real line ($\overline{\Re} \equiv \Re \cup \{-\infty, \infty\}$).[3] I restrict $U_t$ to focus consideration on a special case which is both economically interesting and analytically tractable.

In particular, I assume

$$U_t(c_0, c_1, \ldots, c_t, \ldots) = u(c_t) + \beta \sum_{i=1}^{\infty} \delta^i u(c_{t+i}) \tag{3}$$

where $\delta$ and $\beta$ are discount parameters, and $u(c)$ is a continuous, strictly concave function, $u : [0, \infty) \rightarrow \overline{\Re}$. Unless otherwise stated, I assume that $u$ is a member of the class of CRRA utility functions (with relative risk aversion coefficient $\rho \in (0, \infty)$). When I work in this class, I need $u(c)$ to be well-defined for all $c \in [0, \infty)$. To preserve the continuity property, set $u(0) = \lim_{c \to 0} u(c)$. Hence, for $\rho \in [1, \infty)$, $u(0) = -\infty$.

There are two reasons, other than analytical tractibility, to focus on the preferences in equation 3. The first motivation is that if $u$ is in the CRRA class and $\beta = 1$, the model reduces to the familiar case of exponential dis-

[2]From the perspective of the time zero self, precommitment would be optimal. However, in the real world precommitment is often not possible, since for most forms of economic activity there do not exist institutional mechanisms for precommitment. Three mutually compatible stories may explain this empirical observation. First, such contracts would be difficult to enforce without creating a costly monitoring structure. Second, in a world of uncertainty, such contracts would be difficult to write down. Specifying all of the potential contingencies is impossible. Finally, precommitment is ethically ambiguous, and, at least in the US, some precommitment contracts are not legally binding. Should a 25 year old self be able to make commitments which the 50 year old self is compelled to follow? Which self has the right to make temporally global decisions. The ethical and legal dimensions of this last question are discussed in Schelling (1983).

[3]It may help to interpret self $t$'s payoff as the expectation at time $t$ of this function.

counting with time-additive homothetic preferences. Hence the $\beta \neq 1$ case may be thought of as a perturbation to the "standard" macro preferences. If we care about robustness we probably want to know what happens when such perturbed preferences are considered.

The second motivation is more complex. There is a large body of evidence that discount functions are closely approximated by hyperbolas (*i.e.* the discount function is approximated by the curve $\frac{1}{\kappa_1+\kappa_2\tau}$). This observation was first made by Herrnstein (1961), in relation to animal behavior experiments. The work was later extended with human subjects (DeVilliers and Herrnstein (1976)). Small amendments have been subsequently proposed to this hyperbolic structure, but the basic shape has not been challenged. The important characteristic of the hyperbolic discount function is that it discounts relatively more heavily than the exponential for events in the near future, but discounts less heavily for events in the distant future. Psychologists, notably Ainslie (1975, 1986, 1992), Prelec (1989), and Loewenstein and Prelec (1992) believe that such hyperbolic discounting may play an important role in generating problems of self-regulation, and may provide a potential explanation for numerous behavioral anomalies. When $0 < \beta < 1$ the discount structure in equation [3] mimics the hyperbolic shape, while maintaining most of the analytical tractibility of the exponential case.

The preferences in [3] were first analyzed by Phelps and Pollak (1968). However, their choice of this structure was motivated in a different way. Their game is one of imperfect intergenerational altruism, so the players are (non-overlapping) generations. I assume that the different players are temporally distinct selves of a single person. The mathematical analysis which follows can be applied to either interpretation.

There is another important contrast between this paper and the work of Phelps and Pollak. They confine their analysis to a subset of the joint strategy space by limiting their analysis to symmetric Markov strategies that are linear with respect to the current asset stock. Specifically, they consider equilibria that are supported by the following symmetric strategy for all selves: *Consume at rate* $\lambda$ *whatever the previous history, (set* $c_t = \lambda A_t$*).* They look for $\lambda$ values that support this as a Nash equilibrium.[4] I consider the full strategy space.

Before proceeding with my analysis of this model, it is useful to introduce the following notation. Let $H_t$ be the set of feasible histories of the consumption game at time $t$. An $h_t \in H_t$ history is a $t+1$-element vector, $(A_0, c_0, c_1, c_2, \ldots, c_{t-1}) \in \Re^t$. Let $H_t^F$ represent the set of feasible histories at time $t$. Let $H^F$ be the set of all feasible histories. Let $A : H^F \rightarrow [0, \infty)$ be the map from feasible histories to asset stocks, such that $A(A_0, c_0, c_1, c_2, \ldots, c_{t-1}) \equiv R^t A_0 - \sum_{i=0}^{t-1} R^{t-i} c_i$. Hence, $A(h_t)$ is the asset stock available to self $t$ after history $h_t$. Represent the pure strategy space of self $t$ as,

$$S_t \equiv \{s_t \mid s_t : H_t^F \rightarrow [0, \infty), \ \ and \ \ 0 \leq s(h_t) \leq A(h_t) \ \forall h_t \in H_t^F\}. \qquad (4)$$

Define the joint strategy space $S \equiv \prod_{t=0}^{\infty} S_t$. Let $S^P$ represent the set of subgame-perfect equilibria of the consumption game. Finally, let $v(s, t, h_{\hat{t}})$ represent the continuation payoff of self $t$, after history $h_{\hat{t}}$, when equilibrium strategy $s$ is played from time $\hat{t}$ forward.

# 3 The exponential discount case: $\beta = 1$.

The analysis in this section assumes $\beta = 1$, which implies preferences that are standard to economists: exponential discounting with time-additive utility.

[4] Actually, all of their equilibria satisfy the stronger condition of subgame perfection.

These preferences were first analyzed by Ramsey (1928). His method of analysis was to assume that behavior would correspond to the precommitment strategy of self zero.

**Definition 1:** *A* **Ramsey equilibrium** *is an element of S which is subgame perfect and which maximizes the continuation payoff to self zero in every possible subgame.*

It is easy to show that if $\beta = 1$, $u$ is in the CRRA class, and $\delta R^{1-\rho} < 1$, then there exists a unique Ramsey equilibrium. This equilibrium is summarized by the following consumption rule for all selves: *Consume at rate* $\lambda^R = 1 - (\delta R^{1-\rho})^{\frac{1}{\rho}}$. (The assumption on technology, $\delta R^{1-\rho} < 1$, is standard in the macroeconomics literature, and the condition is assumed to hold for the rest of the paper.) The goal of this section is to determine whether the Ramsey equilibrium is a reasonable prediction for the game-theoretic model of section two. I will ultimately conclude that it is. However, the path to that conclusion is surprisingly challenging.

**Theorem 1:** *Let u be a CRRA utility function with* $\rho \in (0, 1)$. *Fix* $\beta = 1$. *Then the Ramsey equilibrium is the unique subgame perfect equilibrium.*

**Proof:** Based on Theorem 4, and hence postponed.

So far so good. Unfortunately, the conclusion of Theorem 1 does not hold when $\rho \geq 1$. This is particularly worrisome because it is common to calibrate models with $1 \leq \rho \leq 2$.

**Proposition 1:** *Let $u$ be a CRRA utility function with $\rho \in [1, \infty)$. Fix $\beta = 1$. Let $\{c_t^*\}_{t=0}^{\infty}$ be any feasible consumption path. Then $\{c_t^*\}_{t=0}^{\infty}$ can be supported by a subgame perfect equilibrium.*

Proposition 1 says that anything is possible. At first this appears to be a very strong result, but, it is trivial to prove using a rather perverse weakly dominated strategy.

**Proof:** Consider the following equilibrium strategy for self $t$:

*If nobody else has deviated consume $c_t = c_t^*$. If anybody else has deviated consume $c_t = 0$.*

Recall that if $\rho \in [1, \infty)$, then $c_t = 0$ implies $u(c_t) = -\infty$. If some self $s < t$ expects $c_t$ to equal zero, then self $s$'s payoff will be negative infinity. Since deviations produce payoffs of $-\infty$, no self has a strict incentive to deviate. This argument also applies off of the equilibrium path. □

This proof makes use of an unrealistic punishment structure, (though it is subgame perfect). One might want to rule out equilibria which rely on weakly dominated strategies. It is straightforward to show that the only weakly dominated strategy is for a self to consume nothing, or everything. Hence, it seems reasonable to restrict attention to subgame perfect strategies that satisfy the condition $0 < c_t < A_t$ in every subgame in which $A_t > 0$. However, this is not strong enough to eliminate or even to reduce the multiplicy in Proposition 1. This is because it is possible to generate infinite punishments without setting consumption to zero. Consider the case $\rho \in (1, \infty)$. If all selves consume at rate $\lambda$, (*i.e.* $c_t = \lambda A_t \; \forall t$), where $\lambda \in [1 - (\delta R^{1-\rho})^{\frac{1}{\rho-1}}, 1)$, then all selves have payoff $-\infty$.[5] Recall that infinite punishments make anything possible as a perfect equilibrium.

[5] Such infinitely bad punishments with positive consumption levels can also be generated for the case $\rho = 1$.

The next natural step is to see what happens when we exclude equilibria which are supported by such infinitely bad punishments.

**Definition 2:** *A perfect equilibrium has* **finite payoffs at t,** *if the equilibrium satisfies the following condition: For every $h_t \in H_t$ such that $A_t > 0$, all selves $s \geq t$ receive finite payoffs if no deviations occur from time $t$ forward.*

If $\rho \geq 1$ and if an associated subgame perfect equilibrium has finite payoffs at $t$, $\forall t$, then in every subgame with a positive asset stock, the path of future consumption levels will satisfy the restriction $0 < c_t < A_t$. Hence the finite payoff condition rules out weakly dominated strategies. In addition, the finite payoff condition rules out all strategies that rely on infinite punishments. However, the finite payoff criterion does not meaningfully change the equilibrium set.

**Theorem 2:** *Let $u$ be a CRRA utility function with $\rho \in [1, \infty)$. Fix $\beta = 1$. Let $\{c_t^*\}_{t=0}^{\infty}$ be any feasible consumption path, such that $\sum_{i=0}^{\infty} \delta^i u(c_i^*) > -\infty$. Then, $\{c_t^*\}_{t=0}^{\infty}$ can be supported by a perfect equilibrium with finite payoffs at $t$, $\forall t$.*

**Proof:** Let $U_t^* \equiv \sum_{i=0}^{\infty} \delta^i u(c_{t+i}^*)$. So $U_t^*$ is the payoff to self $t$ on the equilibrium path. The statement of the Theorem assumes, $U_0^* > -\infty$, which implies that $\forall t \geq 0$, $U_t^* > -\infty$. Let $A_t^*$ be the inherited asset stock of self $t$ on the equilibrium path.

First, I'll consider the case $\rho = 1$, (*i.e.* $u(\cdot) = \ln(\cdot)$). Construct the set $\{\lambda_{r,r'} \mid r = 1, \ldots, \infty \quad r' = r, \ldots, \infty\}$, according to the following rules. For each $r$, choose $\lambda_{r,r}$ such that $0 < \lambda_{r,r} < 1$, and $U_r^* \geq \delta f(\lambda_{r,r}) + \frac{1}{1-\delta}\ln(A_r^*) + \frac{\delta}{(1-\delta)^2}\ln(R)$, where $f(\lambda) \equiv \frac{1}{1-\delta}[\ln(\lambda) + \frac{\delta}{1-\delta}\ln(1-\lambda)]$. Such a selection is always possible since $f(\lambda) \to -\infty$ as $\lambda \to 0$. Now, given $\lambda_{r,r'}$, with $r' \geq r$, pick $\lambda_{r,r'+1}$ such that $0 < \lambda_{r,r'+1} < 1$ and $\delta f(\lambda_{r,r'+1}) \leq f(\lambda_{r,r'})$. Use this rule to recursively generate the entire set of $\lambda$ values.

Consider the following strategy:

> *If no previous self has deviated, consume $c_t = c_t^*$. If self $r$ was the first to deviate, and self $r'$ was the last to deviate, consume at rate $\lambda_{r,r'}$.*

. The remaining step is to confirm that this strategy supports a perfect equilibrium. Suppose that the current self is self $t$, and that no previous self has deviated. Then,

$$U_t^* \geq \delta f(\lambda_{t,t}) + \frac{1}{1-\delta}\ln(A_t^*) + \frac{\delta}{(1-\delta)^2}\ln(R)$$

$$= \ln(A_t^*) + \sum_{i=1}^{\infty}\delta^i \ln(A_t^* R^i (1-\lambda_{t,t})^{i-1}\lambda_{t,t})$$

$$\geq \ln(c) + \sum_{i=1}^{\infty}\delta^i \ln((A_t^* - c)R^i(1-\lambda_{t,t})^{i-1}\lambda_{t,t}) \quad \forall c \in [0, A_t^*].$$

Hence, no self has an incentive to deviate after histories in which no previous self has deviated.

Now suppose that the history is such that self $r$ was the first to deviate and self $r'$ was the last to deviate. The current period is $t > r'$. WLOG assume $A_t = 1$, and $R = 1$ (for $A_t \neq 1$, $R \neq 1$, carry the appropriate constants through the following inequalities). If self $t$ follows the equilibrium path, its payoff is

$$\sum_{i=0}^{\infty}\delta^i \ln(A_t(1-\lambda_{r,r'})^i R^i \lambda_{r,r'})$$

$$= f(\lambda_{r,r'}) \geq \delta f(\lambda_{r,r'+1}) \geq \ldots \geq \delta^{t-r'} f(\lambda_{r,t}) \geq \delta f(\lambda_{r,t})$$

$$= \ln(A_t) + \sum_{i=1}^{\infty}\delta^i \ln(A_t(1-\lambda_{r,t})^{i-1}R^i\lambda_{r,t})$$

$$\geq \ln(c) + \sum_{i=1}^{\infty}\delta^i \ln((A_t - c)(1-\lambda_{r,t})^{i-1}R^i\lambda_{r,t}) \quad \forall c \in [0, A_t].$$

Hence, no self has an incentive to deviate after histories in which previously selves have deviated. Finally, note that any constant consumption rate bounded strictly between zero and one implies a finite payoff for all selves. This completes the proof for the case $\rho = 1$.

Now consider the case $\rho > 1$. The general argument used for $\rho = 1$ may also be used for this case. However, two differences arise. First, it is necessary to construct a new set $\{\lambda_{r,r'} \mid r = 1, \ldots, \infty \quad r' = r, \ldots, \infty\}$, which is contingent on the value of $\rho$. Moreover, when the $\lambda$ values are chosen, they must all satisfy the condition $\lambda < 1 - (R^{1-\rho}\delta)^{\frac{1}{\rho-1}}$. This guarantees that the continuation payoffs are finite. □

The proof of Theorem 2 is long, but the idea is very simple. Since $\sum_{i=0}^{\infty} \delta^i u(c_i^*) > -\infty$, it can be shown that for all $t$, $\sum_{i=0}^{\infty} \delta^i u(c_{t+i}^*) > -\infty$. Hence, for each self one can construct a finitely bad punishment that ensures that that self will not want to deviate from the equilibrium path. The challenge is to show that these punishments are credible. This is done with sequences of cascading finite punishments, each one worse than the last.

Theorem 1 establishes that the multiplicity in Proposition 1 does not rely on infinitely bad punishments. One consequence of this Theorem and the previous Proposition is that we may now be very suspect of the concept of subgame perfection. Perfection is obviously not strong enough to ensure that the Ramsey equilibrium is the only equilibrium. The rest of this section considers a menu of three mutually compatible refinements, each of which eliminates the multiplicity and leaves the Ramsey solution as the only subgame-perfect outcome.

## 3.1 Phelps and Pollak's uniqueness result.

As stated earlier, Phelps and Pollak (1968) were the first to consider the game described in section two. They looked for equilibria in which all players follow the same linear, Markov strategy both on and off the equilibrium path: *Consume at rate $\lambda$ whatever the previous history.* They did note, however, that other equilibria might exist. Given the assumption on technology, they show that there exists a unique Nash equilibrium which is supported by their rule. Using that result it is straightforward to show that subject to their restrictions there exists a unique subgame perfect equilibrium. This is the Ramsey equilibrium. Obviously, we may be interested in seeing what happens when we consider more general strategy spaces.

## 3.2 Bounded payoffs.

The second refinement is related to the earlier idea of eliminating infinite punishments. The new approach is to consider what happens when we eliminate equilibria that rely on cascades of ever-worsening punishments. The approch is motivated by the following result.

> **Theorem 3:** *Let $u$ be any bounded, continuous real-valued function $u : [0, \infty) \rightarrow \Re$. Fix $\beta = 1$, and $\delta < 1$. Then every subgame-perfect equilibrium is a Ramsey equilibrium.*

**Proof:** Let $C$ represent the space of bounded, continuous functions, $f : [0, \infty) \rightarrow \Re$. Let

$$V(A) \equiv \left\{ v \mid v = v(s, t, h_t), \text{ for some } s \in S^P, \text{ and } h_t \text{ s.t. } A(h_t) = A \right\}.$$

Let $\underline{v}(A) \equiv \inf V(A)$. So $\underline{v} : [0, \infty) \rightarrow \Re$. The Theorem follows from the following three lemmas.

> **Lemma 1:** *$\underline{v}$ is a fixed point of the functional equation,*
>
> $$(Tf)(A) = \max_c \{u(c) + \delta f((A - c)R)\}.$$

**Proof of Lemma 1:** Fix any $A \in [0,\infty)$. The first step is to show, $\underline{v}(A) \leq \max_c\{u(c) + \delta\underline{v}((A-c)R)\}$. Construct a sequence $\{s^n, t^n, h_t^n\}$ such that $A(h_t^n) = A$ for all $n$, and $\lim_{n\to\infty} v(s^n, t^n, h_t^n) = \underline{v}(A)$. Then,

$$v(s^n, t^n, h_t^n) = \max_c\{u(c) + \delta v(s^n, t^n + 1, \{h_t^n, c\})\},$$

which implies that,

$$v(s^n, t^n, h_t^n) \geq \max_c\{u(c) + \delta\underline{v}((A-c)R)\}.$$

Taking the limit of the LHS yields,

$$\underline{v}(A) \geq \max_c\{u(c) + \delta\underline{v}((A-c)R)\}. \tag{5}$$

The next step is to show, $\underline{v}(A) \leq \max_c\{u(c) + \delta\underline{v}((A-c)R)\}$. For this part of the proof I focus exclusively on subgame-perfect equilibria for which the payoff to self one depends exclusively on $A_1$. I assume that $A_0 = A$. Hence, it is possible to represent the continuation payoff to player one as $v(s, 1, (A-c_0)R)$. If $s \in S^P$ then

$$\underline{v}(A) \leq \max_c\{u(c) + \delta v(s, 1, (A-c)R)\}.$$

Construct $\{s^n\}_{n=1}^{\infty}$, such that $s^n \in S^P$ for all $n$, and $v(s^n, 1, \cdot) - \underline{v}(\cdot) \leq \frac{1}{n}$ for all $n$. (Hence, $v(s^n, 1, (A-c)R)$ uniformly converges to $\underline{v}((A-c)R)$.) So,

$$\underline{v}(A) \leq lim_{n\to\infty} \max_c\{u(c) + \delta v(s^n, 1, (A-c)R)\},$$

which implies that,

$$\underline{v}(A) \leq \max_c\{u(c) + \delta\underline{v}((A-c)R)\}. \tag{6}$$

Combining equations 5 and 6 proves the lemma. □

**Lemma 2:** $\underline{v} \in C$.

**Proof of Lemma 2:** First, note that $\underline{v}$ is bounded since $u$ is bounded and $\delta < 1$. It remains to show that $\underline{v}$ is continuous. Fix any $A \in [0,\infty)$, and any sequence $\{A_n\}_{n=1}^{\infty}$ such that $A_n \in [0, \infty)$ for all $n$, and $\lim_{n\to\infty} A_n = A$. Seek to show $\lim_{n\to\infty} \underline{v}(A_n) = \underline{v}(A)$.

Construct a sequence of subgame-perfect equilibrium strategies, $\{s^n\}_{n=1}^{\infty}$ such that $v(s^n, 0, A_n) - \underline{v}(A_n) \leq \frac{1}{n}$. Construct a second sequence of subgame-perfect equilibrium strategies $\{\hat{s}^n\}_{n=1}^{\infty}$ such that $v(\hat{s}^n, 0, A) - \underline{v}(A) \leq \frac{1}{n}$.

Let $\hat{c}^n(A)$ represent the equilibrium consumption of self 0, according to strategy $\hat{s}^n$. Then,

$$v(s^n, 0, A_n) \geq \underline{v}(A) - [u(\hat{c}^n(A)) - u(\hat{c}^n(A) - (A - A_n))],$$

which implies that,

$$\underline{v}(A_n) + \frac{1}{n} \geq \underline{v}(A) - [u(\hat{c}^n(A)) - u(\hat{c}^n(A) - (A - A_n))].$$

Taking limits of both sides yields,

$$\lim_{n\to\infty} \underline{v}(A_n) \geq \underline{v}(A). \tag{7}$$

Let $c^n(A)$ represent the equilibrium consumption of self 0, according to strategy $s^n$. Then,

$$v(\hat{s}^n, 0, A) \geq \underline{v}(A_n) - [u(c^n(A_n)) - u(c^n(A) + (A - A_n))],$$

which implies that,

$$\underline{v}(A) + \frac{1}{n} \geq \underline{v}(A_n) - [u(c^n(A_n)) - u(c^n(A) + (A - A_n))],$$

Taking limits of both sides yields,

$$\underline{v}(A) \leq \lim_{n\to\infty} \underline{v}(A_n). \tag{8}$$

Combining equations 7 and 8 yields

$$\lim_{n\to\infty} \underline{v}(A_n) = \underline{v}(A). \square$$

**Lemma 3:** *(Stokey et al. (1989), Theorem 4.6) The functional equation,*

$$(Tf)(A) = \max_c \{u(c) + \delta f((A - c)R)\}$$

*has a unique fixed point in $C$.*

Together, Lemmas 1, 2, and 3 imply that the continuation payoffs associated with $\underline{v}$ are equivalent to the Ramsey continuation payoffs. □

Theorem 3 establishes that when $\beta = 1$ and the felicity function is bounded (and continuous), all subgame perfect equilibria are Ramsey equilibria.[6] Note that for each value of $\rho$ the CRRA felicity function is unbounded. This explains why Theorem 3 is consistent with Theorem 2. The goal of this section is to develop an equilibrium refinement for the CRRA consumption game which captures the idea of boundedness. Specifically, I restrict attention to equilibria that have the property that for any given equilibrium the set of "normalized" continuation payoffs associated with that equilibrium is bounded. The following definition makes this idea precise.

**Definition 3:** *Fix a strategy profile, $s$. Let*

$$\Omega(s) \equiv \{\omega \mid \exists \text{ a period } t, \text{ feasible history } h_t, \text{ s.t. } A(h_t) > 0 \text{ and } A(h_t)^{1-\rho} \cdot \omega = v(s, t, h_t)\}.$$ [7]

*We will say that $s$ has* **bounded normalized continuation payoffs** *(BNCP) if $\inf \Omega(s)$ is finite. Finally, if $u$ is an instantaneous utility function, let $P^{BNCP}(u)$ represent the associated set of subgame perfect equilibria which satisfy the BNCP property.*

[6] If $u$ is not strictly concave there may not be a unique Ramsey equilibrium.

[7] If $\rho = 1$, let $\Omega(s)$ be as above but replace $A(h_t)^{1-\rho} \cdot \omega$ with $\frac{1-\delta(1-\beta)}{1-\delta} \ln(A(h_t)) + \omega$.

**Theorem 4:** *Let $u$ be a CRRA utility function. Fix $\beta = 1$. The Ramsey equilibrium is the unique subgame perfect equilibrium with the BNCP property.*

**Proof:** There are three relevant cases: $\rho \in (0,1)$, $\rho = 1$, and $\rho > 1$. I'll prove the Theorem for the first case. The treatment of the remaining two cases uses the same approach.

Assume $\rho \in (0,1)$, and assume that there exists a perfect equilibrium $s$ with the property that $W(s)$ has a well-defined (finite) inf, $\underline{\omega}(s)$. Let $A(h)$ represent the current asset stock when the previous history is $h$. Let $\omega(h_t) \equiv v(s,t,h_t)\frac{1-\rho}{(A(h_t))^{1-\rho}}$. The existence of finite $\underline{\omega}(s)$, implies that there exists a sequence of histories $\{h^n_{t(n)}\}_{n=1}^{\infty}$, such that $\underline{\omega}(s) = \lim_{n\to\infty} \omega(h^n_{t(n)})$. Note that the superscript $n$ signifies the location in the sequence. No structure is imposed on the function $t(n)$.

Consider some history $h_t$. If this history arises, the continuation payoff of self $t$ is given by $\frac{(A(h_t))^{1-\rho}}{1-\rho}\omega(h_t)$. Since $s$ is a perfect equilibrium, this payoff is bounded below by,

$$\frac{c^{1-\rho}}{1-\rho} + \frac{\delta[R(A(h_t)-c)]^{1-\rho}}{1-\rho}\underline{\omega}(s) \quad \forall c \in [0, A(h_t)].$$

Let $\hat{c}$ be the value of $c$ that maximizes that lower bound. Using the first order condition (which is necessary and sufficient for this problem) it is easy to show

$$\hat{c} = \frac{A(h_t)}{1 + (\underline{\omega}(s)\delta R^{1-\rho})^{\frac{1}{\rho}}}$$

Plugging this value of $c$ back into the expression for the lower bound and dividing through by $\frac{(A(h_t))^{1-\rho}}{1-\rho}$ yields the following expression for $\omega(h_t)$.

$$\omega(h_t) \geq \left[1 + (\underline{\omega}(s)\delta R^{1-\rho})^{\frac{1}{\rho}}\right]^{\rho}$$

Replacing $h_t$, with $h^n_{t(n)}$ and taking the limit of the LHS as $n \to \infty$, yields,

$$\underline{\omega}(s) \geq \left[1 + (\underline{\omega}(s)\delta R^{1-\rho})^{\frac{1}{\rho}}\right]^{\rho}.$$

Simplification implies that,

$$\underline{\omega}(s) \geq \left[1 + (\delta R^{1-\rho})^{\frac{1}{\rho}}\right]^{-\rho}.$$

Compare this to the payoff that would be received under the Ramsey equilibrium. It is tedious, but not hard to show that if the current period is $t$, with asset stock $A_t$, the Ramsey payoff to self $t$ is equal to

$$\frac{A_t^{1-\rho}}{1-\rho}\left[1 + (\delta R^{1-\rho})^{\frac{1}{\rho}}\right]^{-\rho},$$

so the normalized Ramsey payoff is

$$\left[1 + (\delta R^{1-\rho})^{\frac{1}{\rho}}\right]^{-\rho}.$$

The last inequality on $\underline{\omega}(s)$ implies that the worst normalized continuation payoff of equilibrium $s$ is at least as good as the normalized Ramsey payoff. But the Ramsey payoff is also the upper bound on $W(s)$. It follows that $W(s)$ is a one point set with all continuation payoffs equal to the normalized Ramsey payoff. □

I am now ready to prove Theorem 1.

**Proof of Theorem 1:** Note that if $\rho \in (0,1)$, then $u(\cdot)$ is bounded from below by 0. Hence, if $\rho \in (0,1)$, then any subgame perfect equilibrium has the BNCP property. By Theorem 4, the Ramsey equilibrium is the unique subgame perfect equilibrium. □

## 3.3 Finite horizons.

Consider the finite-horizon analog to the game described in section 2, with $\beta = 1$. If the horizon is $T$, there are $T+1$ selves, and preferences are given by

$$U_t(c_0, c_1, \ldots, c_T) = \sum_{i=0}^{T-t} \delta^i u(c_{t+i}). \tag{9}$$

In subsection 4.2, I show that for the general case $0 < \beta \leq 1$ the finite horizon game has a unique perfect equilibrium. In addition, I show that as $T \to \infty$ these finite horizon equilibria converge to a perfect equilibrium of the infinite horizon game. For the case $\beta = 1$ this limiting equilibrium corresponds to the Ramsey equilibrium. Hence, finite horizon arguments provide another means of picking out the Ramsey equilibrium.

This completes the analysis when $\beta = 1$. For this case, I am very comfortable settling on the Ramsey equilibrium. However, that is not the central issue. More importantly, it may serve the reader to decide, before proceeding, which arguments in favor of the Ramsey outcome are most persuasive. If these refinements are reasonable you may also want to accept their implications in the case $0 < \beta < 1$.

# 4 Non-exponential discounting: $0 < \beta < 1$.

It is trivial to extend the multiplicity results in Proposition 1 and Theorem 1 to the case $0 < \beta < 1$. Hence, it is interesting to see which of the earlier refinements can eliminate this multiplicity. It turns out that only two of the original three refinements continue to imply uniqueness. These two "successful" refinements are discussed first.

## 4.1 Phelps and Pollak revisited.

Once again, Phelps and Pollak's restriction to the set of symmetric, linear Markov strategy, yields a unique perfect equilibrium. Their equilibrium consumption rate, $\lambda^*$ satisfies the equation,

$$\lambda^* = 1 - \left(\delta R^{1-\rho}[\lambda^*(\beta - 1) + 1]\right)^{\frac{1}{\rho}}. \tag{10}$$

Henceforth, I will refer to their equilibrium strategy as the Phelps-Pollak strategy. Note that this reduces to the Ramsey equilibrium strategy when $\beta = 1$.

## 4.2 Finite horizons revisited.

In this subsection, I show that there is a unique perfect equilibrium in the finite-horizon game, and I consider its limiting properties as the horizon goes to infinity. This analysis formalizes the arguments in subsection 3.3, by proving more general results. All of the results in the current subsection apply to the case $0 < \beta \leq 1$.

I begin by analyzing finite horizon games. Note that the $T$-horizon game has $T+1$ players, and preferences given by,

$$U_t(c_0, c_1, \ldots, c_T) = u(c_t) + \beta \sum_{i=1}^{T} \delta^i u(c_{t+i}) \quad \forall t.$$

**Proposition 2:** *For any T-horizon game, there exists a unique subgame-perfect equilibrium. This equilibrium is Markov perfect.*

**Proof:** Let $s_t^T$ be a point in the strategy space of self $t$ in a game with horizon $T$. So $s_t^T : H_t^F \to [0, \infty)$. Suppose that the T-horizon game has a unique perfect equilibrium. Also suppose that this equilibrium has strategies of the form: $s_t^T(h_t) = \lambda_{T-t} A_t$, for all selves $t \in \{0, 1, \ldots, T\}$. Finally, assume $0 < \lambda_{T-t} < 1$ for all $t \in \{0, 1, 2, \ldots, T-1\}$. Let $V(A, T+1) \equiv \beta\delta \sum_{t=0}^{T} \delta^t u(\lambda_{T-t} A_t)$, where $A_0 = A$, and the rest of the $A_t$ sequence is built up recursively: $A_{t+1} = R(1-\lambda_{T-t})A_t$. It is easy to show $V_A(A, T+1) > 0$ and $V_{AA}(A, T+1) < 0 \; \forall \, A \in (0, \infty)$, and $\lim_{A \to 0} V_A(A, T+1) = \infty$. Now consider the behavior of self 0 in a game with horizon $T+1$. Since there is a unique subgame perfect equilibrium in the subgame that arises after self 0's choice, self 0 chooses a consumption level to maximize, $u(c_0) + V(R(A_0 - c_0), T+1)$, subject to the restriction $0 \leq c_0 \leq A_0$. The properties of $u(\cdot)$ and $V(\cdot, T+1)$ imply that this problem has a unique interior solution. It is easy to show that the chosen consumption level is proportional to, but less than, $A_0$. Hence, there exists a number $\lambda$, $0 < \lambda < 1$ such that $c_0 = \lambda A_0 \; \forall A_0$. Set $\lambda_{T+1} \equiv \lambda$. The proof proceeds by induction. To start the induction, simply observe that $\lambda_0 = 1$. □

Since the unique subgame perfect equilibrium in any finite-horizon game is a sequence of Markov strategies, we can write the equilibrium strategy of self $t$ in a $T$-horizon game as, $c_t(A_t, T)$.

**Proposition 3:** *Consider a T-horizon game. The following "Euler equation" holds on the unique equilibrium path.*

$$u'(c_t) = R\delta u'(c_{t+1}) \left[ \frac{\partial c_{t+1}(A_{t+1}, T)}{\partial A_{t+1}} (\beta - 1) + 1 \right] \qquad (11)$$

**Proof:** Continuing the argument from the proof of Proposition 2, note that in equilibrium the following condition holds for all $t$ :

$$u'(c_t) = RV_A(A_{t+1}, T-t).$$

Note that $V(A_{t+1}, T-t)$ can be reexpressed, $V(A_{t+1}, T-t) =$

$$\beta\delta u(c_{t+1}(A_{t+1}, T)) + \delta V(R(A_{t+1} - c_{t+1}(A_{t+1}, T)), T-(t+1))$$

Taking a partial derivative, yields, $V_A(A_{t+1}, T-t) =$

$$\beta\delta u'(c_{t+1}) \frac{\partial c_{t+1}}{\partial A_{t+1}} + \delta R V_A(R(A_{t+1} - c_{t+1}), T-(t+1)) \left[ 1 - \frac{\partial c_{t+1}}{\partial A_{t+1}} \right].$$

Finally substitute $u'(c_{t+1})$ for $RV_A(A_{t+2}, T-(t+1))$ to get the required result.
□

I can now proceed with the main result of this subsection. Theorem 5 establishes that the finite-horizon equilibrium can be used to pick out the Phelps-Pollak equilibrium in the infinite horizon game.

**Theorem 5:** *As $T \to \infty$, $c_t(A, T)$ converges pointwise to the function $\lambda^* A$, which is the symmetric Markov strategy in the Phelps and Pollak equilibrium.*

**Proof:** Recall that the proof of Proposition 2 shows that in a game with horizon $T$, the unique perfect equilibrium strategy of self $t$ is to consume at rate $\lambda_{T-t}$. Given this observation, it is possible to use Proposition 3 to characterize the consumption of self $t$ in a $T+1$-horizon game.

Recall that Proposition 3 states that the following equation holds on the unique equilibrium path of any finite horizon subgame:

$$u'(c_{t-1}) = R\delta u'(c_t)\left[\frac{\partial c_t(A_t,T)}{\partial A_t}(\beta-1)+1\right]$$

Assume that the game has horizon $T$. Substitute in for $u(\cdot)$, replace $c_t$ with $\lambda_{T-t}A_t$, and replace the partial derivative with $\lambda_{T-t}$. Solving for $c_{t-1}$ yields,

$$c_{t-1} = \lambda_{T-(t-1)}A_{t-1},$$

where

$$\lambda_{T-(t-1)} = \frac{\lambda_{T-t}}{[\delta R^{1-\rho}(\lambda_{T-t}(\beta-1)+1)]^{\frac{1}{\rho}} + \lambda_{T-t}} \tag{12}$$

This implies that in a finite horizon game it is possible to calculate the equilibrium consumption rate of today's self from the equilibrium consumption rate of tomorrow's self. Another way of thinking about this is to say that it is possible to calculate self $t$'s equilibrium strategy in the $T+1$-horizon game if we know self $t$'s equilibrium strategy in the $T$-horizon game.

So far I've noted the following properties. First $c_t(A_t,T) = \lambda_{T-t}A_t$ both on and off the equilibrium path. Second, the sequence of consumption rates $\{\lambda_\tau\}_{\tau=0}^{\infty}$, follows the recursion,

$$\lambda_{\tau+1} = f(\lambda_\tau) \equiv \frac{\lambda_\tau}{[\delta R^{1-\rho}(\lambda_\tau(\beta-1)+1)]^{\frac{1}{\rho}} + \lambda_\tau}$$

Hence, to prove the Theorem it is sufficient to show that $\lambda_\tau \to \lambda^*$.

In the argument which follows I'll use the following properties of $f(\cdot)$, which are straightforward to confirm.

- $f(0) = 0$.
- $f(\cdot)$ differentiable on $[0,1]$.
- $f'(0) > 1$ (using technology assumption $\delta R^{(1-\rho)} < 1$).
- $f'(x) > 0$ on $[0,1]$.
- $f(1) < 1$.

Let $\overline{\lambda} = \sup\{\lambda \mid \lambda \in [0,1], \lambda = f(\lambda)\}$. There is at least one fixed point at zero, so $\overline{\lambda}$ exists. In fact, it is possible to show that $\overline{\lambda}$ is strictly greater than zero. This follows from the properties $f(0) = 0$, $f'(0) > 1$, $f(\cdot)$ continuous, $f(1) < 1$, and by application of the Intermediate Value Theorem. Finally, $\lambda_r \to \overline{\lambda}$ since $f'(x) > 0$ on $[0,1]$, $f(1) < 1$, and $\lambda_0 = 1$.

It only remains to show that $\overline{\lambda} = \lambda^*$. Recall that $\overline{\lambda} = f(\overline{\lambda})$. Both sides of this equation can be divided by $\overline{\lambda}$ since it has been shown that $\overline{\lambda} > 0$. Transforming the resulting equality it is easy to show that

$$\overline{\lambda} = 1 - (\delta R^{1-\rho})^{\frac{1}{\rho}} \left[\overline{\lambda}(\beta - 1) + 1\right]^{\frac{1}{\rho}}.$$

Note that the unique solution to this equation is $\lambda^*$. □

Before proceeding with the next subsection, it is helpful to extend the analysis of this subsection a little further. Recall equation 11, and note that Theorem 5 establishes that as the horizon goes to infinity

$$\frac{\partial c_{t+1}(A_{t+1}, T)}{\partial A_{t+1}} \longrightarrow \lambda^*$$

Hence, in the limit,[8] equation 11 becomes,

$$u'(c_t) = R\delta u'(c_{t+1}) \left[\lambda^*(\beta - 1) + 1\right] \qquad \forall t \tag{13}$$

Note that this equation and its finite-horizon analog reduce to the standard Euler equation when $\beta = 1$. It is also interesting to observe that equation 13 is observationally equivalent to the Euler equation that arises when discounting is exponential with discount rate $\overline{\delta} = \delta\left[\lambda^*(\beta - 1) + 1\right]$. This is an observational equivalence result which is somewhat similar to the widely

[8]This can be interpreted formally in the following way. Consider the infinite horizon game. Let $s^*$ be the equilibrium of that game which is picked out in Theorem 5. Equation 13 holds on the equilibrium path of $s^*$.

cited but false observational equivalence claim of Strotz (1955).[9] Note that my derivation of equation 13 depends on two assumptions that Strotz did not make. First, my discounting structure is very restricted, reverting to exponential discounting after one period. Second, I assume CRRA preferences. The specific form of my result depends on both of these assumptions.[10]

However, there is a sense in which my observational equivalence claim is just a special case of a more general phenomenon. With CRRA preferences and a fixed interest rate, any constant consumption rate implies a linear relationship between $u'(c_t)$ and $u'(c_{t+1})$. If the consumption rate is $\lambda$,

$$u'(c_t) = R\hat{\delta}u'(c_{t+1})$$

where $\hat{\delta} = R^{\rho-1}(1-\lambda)^{\rho}$.

## 4.3 An interesting class of equilibria.

I now return to the general infinite-horizon case. Before discussing the remaining two refinements, it is helpful to describe a class of equilibria to which I will later refer. Consider the following symmetric strategy (symmetric in the sense that all selves follow the same consumption rule).

> *Consume a fraction $\lambda_0 \leq \lambda^*$ of the current asset stock unless some prior self has consumed at a rate different than $\lambda_0$. In that case, consume at rate $\lambda^*$.*

[9] Strotz claimed that dynamically inconsistent agents would behave as if they had an exponential discount rate equal to their instantaneous rate of discount at time zero. See Phelps (1968) for an explanation of what Strotz did wrong.

[10] With completely general discounting a very different "Euler-equation" is generated. This general equation contains an infinite sequence of marginal utilities and propensities to consume.

Call this strategy the Self-Diagnostic rule ($SD$ rule). The rule specifies that selves choose a (cooperative) low consumption rate, unless some previous self has violated the rule. In the case of a previous violation, the current self is instructed to consume at the Phelps-Pollak rate. The strategy uses linear consumption rules both on and off the equilibrium path, and admits the Phelps-Pollak strategy as a special case (*i.e.* $\lambda_0 = \lambda^*$). Also note that the equilibrium path is supported by a "focal" strategy, in the sense that the post-deviation phase corresponds to the unique equilibrium picked out in the previous two subsections. Since the $SD$ strategy depends on both $\lambda_0$ and $\lambda^*$, represent it as $SD_{\lambda_0,\lambda^*}$.

**Proposition 4:** *Fix $0 < \beta < 1$, and $0 < \delta < 1$. There exists an $\epsilon > 0$ such that for all $\lambda_0$ in the interval $(\lambda^* - \epsilon, \lambda^*)$, the $SD_{\lambda_0,\lambda^*}$ rule supports a subgame perfect equilibrium which Pareto-dominates the outcome associated with the Phelps-Pollak equilibrium.*

Note that two equilibria are Pareto-rankable if all temporal selves of the individual are better off under one of the two equilibria. Proposition 4 is proved by extending a result in Phelps and Pollak (1968).

**Proof:** Phelps and Pollak show that for small $\epsilon$ there exists an interval $(\lambda^* - \epsilon, \lambda^*)$, with the property that if $\lambda$ is in that interval, an infinite path of consumption at rate $\lambda$ Pareto-dominates the consumption path of the Phelps-Pollak equilibrium.

The remaining step is to show that this Pareto-superior path is supported as a perfect equilibrium by the $SD$ rule. Assume that at time $t$ no previous self has deviated. So the $SD$ rule implies that self $t$ should consume at rate $\lambda \in (\lambda^* - \epsilon, \lambda^*)$. Suppose self $t$ were to deviate. The $SD$ rule dictates that a current deviation is punished by future consumption at rate $\lambda^*$. So what is self $t$'s best deviation? Recall that $\lambda^*$ has the property that if all future selves consume at $\lambda^*$ then the current self will also want to consume at $\lambda^*$. So self $t$'s best deviation is to consume at rate $\lambda^*$. However, we know that consumption at rate $\lambda^*$ forever makes self t no better off than consumption at rate $\lambda$ forever. So there is no incentive to deviate from the equilibrium path. By definition of $\lambda^*$ there is also no incentive to deviate after a history which is off of the equilibrium path. □

One useful implication of this proof is that if an $SD_{\lambda_0,\lambda^*}$ strategy supports a perfect equilibrium than the corresponding equilibrium path Pareto-dominates the equilibrium path of the Phelps-Pollak equilibrium.

## 4.4 The BNCP refinement.

Recall the notation of section 3.2. It is trivial that if $s$ is an $SD$ equilibrium, $\Omega(s)$ is a two point set. It can also be shown that these two points are finite. Hence, every $SD$ equilibrium is admissible by the BNCP criterion. Since the $SD$ equilibria are indexed by a nondegenerate interval the BNCP criterion is too weak to provide uniqueness.

## 4.5 Which equilibrium is focal when $0 < \beta < 1$?

Having considered the three refinements, it is now possible to ask if there is a focal equilibrium for the case $0 < \beta < 1$. The most likely candidate is the Phelps-Pollak equilibrium. Subsections 4.1 and 4.2 showed that this equilibrium is uniquely chosen by the Phelps-Pollak refinement and the finite horizon refinement. In addition, the Phelps-Pollak equilibrium has appealingly simple symmetric strategies which satisfy the Markov property that the sufficient statistic for today's consumption decision is today's asset stock. Finally, the Phelps-Pollak equilibrium satisfies the BNCP criterion.

However, there is also a strong argument against the Phelps-Pollak equilibrium. Proposition 4 shows that there exist $SD$ equilibria that Pareto-dominate the Phelps-Pollak equilibrium. Moreover, the $SD$ rules are symmetric and relatively simple (though they do not have the Markov property), and the $SD$ equilibria satisfy the BNCP criterion.

Finally, there are other perfect equilibria outside of the $SD$ class which also Pareto-dominate the Phelps-Pollak equilibrium. Is the Phelps-Pollak outcome, an $SD$ outcome, or some other outcome most likely? When $\beta = 1$ we had a clear answer. For the case $0 < \beta < 1$, there is not an obvious choice.[11]

# 5 Characterization of the equilibrium set.

This section characterizes the set of subgame perfect equilibria which survive the BNCP criterion, when $u$ is a CRRA utility function. Hence the section characterizes the set $P^{BNCP}(u)$, which I will henceforth shorten to $P^{BNCP}$. The characterization depends upon two assumptions (in addition to the earlier technology assumption) which respectively restrict the range of $\delta$ and $\beta$, the discount parameters in my model. The restriction on $\delta$ takes the form,

$$\text{(A1)} \qquad (\delta R^{1-\rho})^{\frac{1}{\rho}} > \tfrac{1}{2}$$

This inequality is satisfied for sufficiently large $\delta$. The restriction on $\beta$ is less clear-cut. In particular, I assume,

(A2) $\beta$ sufficiently close to 1.

Recall that $\beta = 1$ is the standard exponential discount case. When the model is calibrated with reasonable parameter values, assumption (A2) is

[11] I am currently working on renegotiation refinements. Preliminary work in this area indicates that some renegotiation criteria, (*e.g.*, Farrell and Maskin's (1989) weak renegotiation-proofness criterion), continue to admit multiple equilibria.

not restrictive – *i.e.* the results which depend on (A2) hold for $\beta$ values in an interval $(\underline{\beta}, 1]$, where $\underline{\beta}$ is "close" to zero. At the end of this section I present some examples which support this claim.

Consider the set, $P^{BNCP}$, which contains all of the subgame perfect strategy profiles of the BNCP class. Let $\underline{\omega}$ represent the worst normalized continuation payoff in the set of all normalized continuation payoffs of strategy profiles in $P^{BNCP}$.

$$\underline{\omega} \equiv \min\{\, \omega \,|\, \exists\, s \in P^{BNCP}, \text{ s.t. } \omega \in \Omega(s)\}.$$

For now, I'll assume that $\underline{\omega}$ is well-defined. A formal existence result appears in the main theorem.

Before proceeding to the theorem, it may be helpful to emphasize the reason that we care about $\underline{\omega}$. Let,

$$\omega(\{\lambda_n\}_{n=m}^{\infty}) \equiv u(\lambda_m) + \beta \sum_{i=1}^{\infty} \delta^i u(R^i \lambda_{m+i} \prod_{j=0}^{i-1} (1 - \lambda_{m+j})),$$

so $\omega(\{\lambda_n\}_{n=m}^{\infty})$ is the normalized payoff to the current self if the path of current and future consumption rates is given by $\{\lambda_n\}_{n=m}^{\infty}$.

**Proposition 5:** *A path of consumption rates, $\{\lambda_t\}_{t=0}^{\infty}$ can be supported by a perfect equilibrium with bounded normalized continuation payoffs iff $\underline{\omega} \leq \omega(\{\lambda_n\}_{n=m}^{\infty})\ \forall\ m \geq 0$.*

**Proof:** Necessity follows immediately from the definition of $\underline{\omega}$. To show sufficiency, it is helpful to introduce the following notation. Let $\lambda(s, t, h)$ represent the equilibrium consumption rate of self $t$ after observing history $h$ when the equilibrium strategy profile is $s$. Pick $s^* \in P^{BNCP}$, $t^*$, and $h^*$, such that [12]

$$\underline{\omega} = \frac{v(s^*, t^*, h^*)}{A(h^*)^{1-\rho}}.$$

Note that such a triplet must exist by the definition of $\underline{\omega}$. Finally, for any two time periods $r$ and $t$, with $r < t$, let $h(r,t) \equiv \{h^*, \hat{\lambda}_r, \hat{\lambda}_{r+1}, \ldots, \hat{\lambda}_{t-1}\}$, where the respective $\hat{\lambda}$'s are the realized sequence of consumption rates from time $r$ to time $t-1$, and $h^*$, is the history of consumption rates defined above. Now take a path of consumption rates, $\{\lambda_t\}_{t=0}^{\infty}$, such that $\underline{\omega} \leq \omega(\{\lambda_n\}_{n=m}^{\infty})$ $\forall\, m \geq 0$. Consider the following strategy profile:

> *If no previous self has deviated, self $t$ consumes at rate $\lambda_t$. If $r < t$ was the first self to deviate, self $t$ consumes at rate $\lambda(s^*, t^* + t - r, h(r,t))$.*

This profile is a subgame perfect equilibrium which supports the proposed path of consumption rates. □

The main theorem depends on the property that the Ramsey consumption path, (*i.e.* $\lambda_t = \lambda^R \equiv 1 - (\delta R^{1-\rho})^{1/\rho}$ $\forall t$), can be supported by a subgame perfect strategy profile with bounded normalized continuation payoffs. Proposition 6 provides a sufficient condition for this property.

**Proposition 6:** *Given (A1) and (A2), the Ramsey consumption path can be supported by a subgame perfect strategy profile with bounded normalized continuation payoffs.*

**Proof:** Let $\omega(\lambda) \equiv \omega(\lambda, \lambda, \lambda, \ldots)$. Let $f(\beta) \equiv \omega(\lambda^R) - \omega(\lambda^*(\beta))$, where $\lambda^*(\beta)$ is the Phelps and Pollak consumption rate. This notation is used to emphasize that $\omega(\lambda^*(\beta))$ depends on $\beta$ in two ways. First, $\beta$ is a discount rate, so changes in $\beta$ affect the value of future consumption. Second, $\beta$ is in

[12] When $\rho = 1$, find $s^* \in P^{BNCP}$, $t^*$, and $h^*$, such that $\underline{\omega} = v(s^*, t^*, h^*) - \frac{1-\delta(1-\beta)}{1-\delta} \ln(A(h^*))$.

the implicit equation which determines $\lambda^*$. Note that $\beta$ only influences $\omega(\lambda^R)$ through the former mechanism as $\lambda^R$ does not depend on $\beta$.

The body of this proof characterizes the value of $f(\cdot)$ in a neighborhood of $\beta = 1$. First, $f(1) = 0$ since $\lambda^*(1) = \lambda^R$. To evaluate $f(\cdot)$ at $\beta$ values just below unity, I consider $f'(1)$ and $f''(1)$.

$$f'(\beta) = \frac{\partial\omega(\lambda^R)}{\partial\beta} - \frac{\partial\omega(\lambda^*)}{\partial\beta} - \frac{\partial\omega(\lambda^*)}{\partial\lambda^*}\frac{d\lambda^*}{d\beta}$$

Note that $f'(1) = 0$, as $\frac{\partial\omega(\lambda^R)}{\partial\beta}\Big|_{\beta=1} = \frac{\partial\omega(\lambda^*)}{\partial\beta}\Big|_{\beta=1}$, and, $\frac{\partial\omega(\lambda^*)}{\partial\lambda^*}\Big|_{\beta=1} = \frac{\partial\omega(\lambda^R)}{\partial\lambda^R}\Big|_{\beta=1} = 0$. Tedious algebraic manipulations reveal that

$$f''(1) = \frac{1}{\rho}(\lambda^R)^{-\rho}(1-\lambda^R)(1-2\lambda^R),$$

which is positive by (A1) (recall that $\lambda^R = 1 - (\delta R^{1-\rho})^{1/\rho}$). Given that $f(1) = 0$, $f'(1) = 0$, and $f''(1) > 0$, there exists an interval $(\underline{\beta}, 1)$ such that $f(\beta) > 0 \;\forall\; \beta \in (\underline{\beta}, 1)$. Hence, for sufficiently large $\beta < 1$, and $\delta$ satisfying (A2), $f(\beta)$ is positive and the $SD_{\lambda^R,\lambda^*}$ rule is a subgame perfect equilibrium. □

**Theorem 5:** *Assume that the Ramsey consumption path can be supported by a subgame perfect strategy profile with bounded normalized continuation payoffs. Then $\underline{\omega}$ is well defined, and*

$$\underline{\omega} = \omega(\lambda^D(\overline{\lambda}^P), \overline{\lambda}^P, \lambda^R, \ldots) = \omega(\overline{\lambda}^P, \lambda^R, \ldots),$$

*where,*

$$\lambda^D(\lambda^P) \equiv \underset{\lambda\in[0,1]}{\operatorname{argmax}}\; \omega(\lambda, \lambda^P, \lambda^R, \ldots),$$

$$\overline{\lambda}^P \equiv \max\{\lambda^P \in (0,1) \,|\, \omega(\lambda^D(\lambda^P), \lambda^P, \lambda^R, \ldots) = \omega(\lambda^P, \lambda^R, \ldots)\}.$$

**Proof:** First, note that $\lambda^D(\lambda^P)$ is a function on $(0,1)$. Let

$$g(\lambda^P) \equiv \omega(\lambda^P, \lambda^R, \ldots) - \omega(\lambda^D(\lambda^P), \lambda^P, \lambda^R, \ldots).$$

Let $\lambda^M \equiv \text{argmax}_{\lambda \in [0,1]} \omega(\lambda, \lambda^R, \ldots)$. The following results are straightforward to confirm:

- $\lambda^M$ is a point in $(0,1)$,
- $\frac{\partial^2 g(\lambda^P)}{(\partial \lambda^P)^2} < 0$ for all $\lambda^P > \lambda^M$,
- $g(\lambda^M) > 0$,
- $\lim_{\lambda^P \to 1} g(\lambda^P) < 0$.

Hence, $\overline{\lambda}^P$ is well-defined. The rest of the proof is based on the following result.

**Lemma 4:** *Given $\underline{\omega}$, let $\overline{\lambda} \equiv sup\{\lambda \mid \omega(\lambda, \lambda^R, \ldots) = \underline{\omega}\}$. Then,*

$$\underline{\omega} = \omega(\overline{\lambda}, \lambda^R, \ldots) = \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots).$$

**Proof of Lemma 4:** Fix any subgame perfect equilibrium, $s$, with normalized continuation payoff to self 0 of $\omega$. Then,

$$\omega \geq u(\lambda) + \delta[(1-\lambda)R]^{1-\rho}[\omega(s,1,\lambda) - (1-\beta)u(\lambda_1(\lambda))], \quad \text{for all } \lambda \in [0,1],$$

where $\omega(s,1,\lambda)$ represents the normalized payoff to self 1 under equilibrium $s$ if self 0 has played $\lambda$; and $\lambda_1(\lambda)$, represents the equilibrium map from self 0's action $\lambda$, to self 1's action, $\lambda_1$. By assumption, $\underline{\omega} \leq \omega(s,1,\lambda)$, for all $s, \lambda$, so,

$$\omega \geq u(\lambda) + \delta[(1-\lambda)R]^{1-\rho}[\underline{\omega} - (1-\beta)u(\lambda_1(\lambda))], \quad \text{for all } \lambda \in [0,1].$$

It is straightforward to confirm that

$$\overline{\lambda} = \sup\{\lambda \mid \exists \{\lambda_n\}_{n=1}^{\infty} \text{ such that}, \omega(\lambda, \{\lambda_n\}_{n=1}^{\infty}) \geq \underline{\omega}\}.$$

Hence, $\lambda_1(\lambda) \leq \overline{\lambda}$, implying,

$$\omega \geq u(\lambda) + \delta[(1-\lambda)R]^{1-\rho}[\underline{\omega} - (1-\beta)u(\overline{\lambda})], \text{ for all } \lambda \in [0,1],$$

which in turn implies,

$$\omega \geq \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots).$$

By definition of $\underline{\omega}$, there exists a sequence of $\omega$'s which converge to $\underline{\omega}$, and for each $\omega$ in this sequence the previous inequality holds. Hence,

$$\underline{\omega} \geq \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots). \quad (14)$$

Construct a monotonically increasing sequence, $\{\lambda_n\}_{n=1}^{\infty}$ which converges to $\overline{\lambda}$. By definition of $\overline{\lambda}$ the following *equilibrium path* strategy is supportable with a subgame-perfect equilibrium:

*Self 0: consume at rate* $\lambda^D(\lambda^n)$.
*Self 1: consume at rate* $\lambda^n$.
*Selves* $t \geq 2$: *consume at rate* $\lambda^R$.

Hence $\underline{\omega} \leq \omega(\lambda^D(\lambda^n), \lambda^n, \lambda^R, \ldots)$. Letting $\lambda^n$ go to unity (and noting that $\lambda^D$ is continuous in its argument), yields

$$\underline{\omega} \leq \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots). \quad (15)$$

Combining equations 14 and 15, yields,

$$\underline{\omega} = \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots),$$

proving the lemma. □

To complete the proof of Theorem 5 it is sufficient to show that given $\underline{\omega}$, $\overline{\lambda} = \overline{\lambda}^P$. Suppose $\overline{\lambda} > \overline{\lambda}^P$. Then, by definition of $\overline{\lambda}^P$,

$$\omega(\overline{\lambda}, \lambda^R, \ldots) \neq \omega(\lambda^D(\overline{\lambda}), \overline{\lambda}, \lambda^R, \ldots),$$

contradicting Lemma 4. Alternatively, assume, $\overline{\lambda} < \overline{\lambda}^P$. By definition of $\overline{\lambda}^P$, the following ***equilibrium path*** strategy is supportable by a subgame-perfect equilibrium:

*Self 0: consume at rate* $\lambda^D(\overline{\lambda}^P)$.
*Self 1: consume at rate* $\overline{\lambda}^P$.
*Selves* $t \geq 2$*: consume at rate consume at rate* $\lambda^R$.

Hence, by definition of $\overline{\lambda}^P$ and $\overline{\lambda}$ the normalized payoff to self 0 is

$$\omega(\lambda^D(\overline{\lambda}^P), \overline{\lambda}^P, \lambda^R, \ldots) = \omega(\overline{\lambda}^P, \lambda^R, \ldots) < \omega(\overline{\lambda}, \lambda^R, \ldots) = \underline{\omega},$$

contradicting the definition of $\underline{\omega}$. Hence, $\overline{\lambda} = \overline{\lambda}^P$. □

Theorem 5 characterizes the worst continuation payoff which can be supported by a perfect equilibrium of the BNCP class. Proposition 7, which follows, characterizes the best continuation payoff which can be supported by a perfect equilibrium of the BNCP class. Let,

$$\overline{\omega} \equiv \max\{\,\omega \,|\, \exists\, s \in P^{BNCP}, \text{ s.t. } \omega \in \Omega(s)\}.$$

**Proposition 7:** ***Assume that the Ramsey consumption path can be supported by a subgame perfect strategy profile with bounded normalized continuation payoffs. Then $\overline{\omega}$ is equal to the normalized payoff which would be achieved by self zero if self zero could precommit all future consumption rates.***

**Proof:** It is sufficient to note that from the perspective of self zero the optimal path of consumption rates from period one forward is given by $\lambda_t = \lambda^R$. □

Theorem 5 and Proposition 7 rely on an assumed property: there exists an element of $P^{BNCP}$ which supports the Ramsey consumption path. It is useful to know if this property holds when the model is calibrated with standard parameter values. Proposition 6 establishes that assumptions (A1) and (A2) are sufficient for this property to hold. Hence, an interesting exercise is to determine the restrictiveness of (A1) and (A2). Do these assumptions rule out any of the parameter values we would like to use to calibrate this model?

Condition (A1) can be analyzed directly. Table 1 examines three leading cases which span the range of $\rho$ values from which most consumption models are calibrated. Inspection of the table immediately reveals that (A1) is not restrictive since preferences are usually calibrated with $1 \leq \rho \leq 2$, and $.90 \leq \delta \leq 1$.

Condition (A2) cannot be analyzed as easily. Recall, that I seek to characterize the calibration range over which there exists a subgame perfect equilibrium which implements the Ramsey consumption path. A sufficient condition for the existence of such an equilibrium, is the existence of an $SD_{\lambda^R,\lambda^*}$ equilibrium. This is the approach taken in the proof of Proposition 6. Table 2 extends this analysis by mapping $\rho, \delta$ pairs into intervals of $\beta$ values which have the property that if $\beta$ is in that interval a $SD_{\lambda^R,\lambda^*}$ equilibrium exists.

Table 1: Restrictiveness of assumption (A1)

| | (A1) given $\rho$ from column 1 | (A1) given $\rho$ and given $R = 1.03$ |
|---|---|---|
| $\rho = \frac{1}{2}$ | $\delta^2 R > \frac{1}{2}$ | $\delta > .70$ |
| $\rho = 1$ | $\delta > \frac{1}{2}$ | $\delta > .50$ |
| $\rho = 2$ | $\sqrt{\frac{\delta}{R}} > \frac{1}{2}$ | $\delta > .26$ |

Table 2: Restrictiveness of assumption (A2)

| $\beta$ intervals for which $\exists\ SD_{\lambda_R,\lambda_*}$ equilibrium given $R = 1.03$, and various $\rho, \delta$ pairs | | |
|---|---|---|
| | $\delta = .9$ | $\delta = .95$ |
| $\rho = \frac{1}{2}$ | $.334 < \beta \leq 1$ | $.245 < \beta \leq 1$ |
| $\rho = 1$ | $.003 < \beta \leq 1$ | $.012 < \beta \leq 1$ |
| $\rho = 2$ | $0 < \beta \leq 1$ | $0 < \beta \leq 1$ |

Specifically, Table 2 reports the intervals of $\beta$ values for which

$$\omega(\lambda^R, \ldots) \geq \omega(\lambda^*, \ldots).$$

Table 2 demonstrates that (A2) is not restrictive. For each of the examples that I consider there is a large range of $\beta$ values for which the Ramsey path can be implemented. Moreover, for $\rho = 2$ the entire unit interval (open at zero) is in the acceptable region.

# 6 Interpreting the results.

In this section I argue that the preceding multiplicity is actually a good thing, since it provides a potential explanation for some puzzling economic phenomena. I consider the consequences of this analysis for the theory of dynamic choice, and then turn to some specific applications which may be of interest to macroeconomists.

## 6.1 Dynamic choice.

There are three general conclusions which I want to highlight. First, since there is not a focal equilibrium in the case $0 < \beta < 1$, it seems reasonable to predict that two people with identical preferences would exhibit different behavior in the same environment. Hence, the model generates heterogeneous behavior without making recourse to heterogeneous preferences.

Second, this theory explains why some people have self-acknowledged "bad habits" which they can't break, while other people with identical preferences are able to exert what might be called "self-control" to avoid these "bad habits." The bad habit is a perfect equilibrium which is Pareto-dominated by another perfect equilibrium. Consider a person whose selves are playing the Phelps-Pollak equilibrium, and contrast this individual with a different person whose selves are consuming at a lower rate because they are playing a Pareto-superior equilibrium of the *SD* class. The Phelps-Pollak individual may wish that she had the "self-control" of the *SD* individual. However, the Phelps-Pollak individual is in equilibrium, and, short of some kind of renegotiation, will not achieve the better outcomes of the *SD* person.

Third, the theory makes sense of the observation that behavior is often perceived to be self-diagnostic. On the equilibrium path of the *SD* equilibrium, one self-indulgent act—a high consumption rate today—begets a sequence of self-indulgent acts—high consumption rates in the future.

## 6.2 Macroeconomic applications.

The model is of direct interest to macroeconomists, since the intertemporal choice problem that I study is a savings/consumption decision. Most importantly, the model explains how an economy's/individual's savings rate can be indeterminate in equilibrium, without using a traditional externality argument. The scope of this indeterminacy is demonstrated in the following example.

Let $\underline{\lambda}$ be the smallest $\lambda_0$ value for which the $SD_{\lambda_0,\lambda^*}$ strategy supports a perfect equilibrium. It can be shown that for all $\lambda \in [\underline{\lambda}, \lambda^*]$, the $SD_{\lambda,\lambda^*}$ rule supports a perfect equilibrium. Table 3 provides examples of the $[\underline{\lambda}, \lambda^*]$ interval for a small set of $\beta, \delta$ pairs when $\rho = 1$.[13]

---

[13]It can be shown that $\rho = 1$ implies that the interval is independent of $R$. $\lambda^* = \frac{1-\delta}{1-\delta(1-\beta)}$, and $\underline{\lambda}$ is the smaller of the two solutions to the following non-linear equation:

$$(\lambda^*)^{-1} \ln\left(\frac{\lambda}{\lambda^*}\right) + \frac{\beta\delta}{(1-\delta)^2} \ln\left(\frac{1-\lambda}{1-\lambda^*}\right) = 0.$$

**Table 3:** $[\underline{\lambda}, \lambda^*]$

| | $\beta = .25$ | $\beta = .50$ | $\beta = .75$ | $\beta = 1.00$ |
|---|---|---|---|---|
| $\delta = .975$ | [.003,.093] | [.011,.049] | [.019,.033] | [.025,.025] |
| $\delta = .950$ | [.009,.174] | [.024,.095] | [.038,.066] | [.050,.050] |
| $\delta = .925$ | [.017,.245] | [.040,.140] | [.059,.098] | [.075,.075] |
| $\delta = .900$ | [.030,.308] | [.057,.182] | [.080,.129] | [.100,.100] |

Consider the following particular example. If $\beta = .50$, $\delta = .975$, and an $SD$ equilibrium is adopted, then the individual's long-run consumption rate can be as low as .011 and as high as .049. It is helpful to contrast two hypothetical people (who may be interpreted as two small open economies). Suppose the first person is in an $SD$ equilibrium with $\lambda = .011$. Suppose the second person is in an $SD$ equilibrium with $\lambda = .049$. Take the interest rate to be 3%, (*i.e.* R=1.03). Then the consumption of individual one will grow exponentially over time, while the consumption of individual two will fall exponentially; individual one exhibits a savings rate of 62.2% of income, while individual two exhibits a savings rate of -67.8% of income. However, in terms of their deep preferences these two people are completely identical. They function in the same institutional environment. The only difference is that they implement different perfect equilibria. Moreover, it is hard to argue in favor (from a positive perspective) of one equilibrium over the other. The high consumption rate equilibrium is the Phelps-Pollak equilibrium, and has all of the nice properties discussed above. However, it is Pareto-dominated by the low consumption rate equilibrium. Which equilibrium is more likely to occur?

This analysis suggests a potential rationale for government intervention to try to raise the national savings rate. I have analyzed a model in which low savings rate are Pareto-inferior, but nevertheless arise as an equilibrium outcome. Hence it may make sense to use social institutions to pick out welfare-enhancing equilibria. Perhaps schools should teach students to save?

Finally, the analysis of this paper can be interpreted as a general methodological critique of the "Euler equation" approach to consumption. When $\beta$ is less than one there are multiple equilibria, and hence, the usual consumption Euler equation no longer holds. (If a unique Euler equation did hold than there could not be multiple equilibria.) Moreover, on a generic subgame-perfect equilibrium path there will be no systematic relationship between the interest rate and the path of marginal utilities. Marginal utilities take on very little importance because for generic equilibria the decision to deviate is not based on considerations of local perturbations to the consumption path.[14]

# 7 Conclusion.

This paper characterizes the equilibria that arise when dynamic savings decisions are modelled as an intra-personal game. I present three refinements which admit a unique equilibrium (the Ramsey equilibrium) when discounting is exponential. However, only two of these refinements continue to admit a unique equilibrium when discounting is approximately hyperbolic. Moreover, the unique equilibrium in those two "successful" cases turns out to be Pareto-inferior in the class of subgame perfect equilibria. This leads me to conclude that there is no single focal equilibrium in the case of hyperbolic discounting. I consider several consequences of this result, with emphasis on the indeterminacy of the national savings rate.

There are three extensions on which I am currently working. The first is to incorporate renegotiation refinements. Preliminary work in this area indicates that some renegotiation criteria, (*e.g.* Farrell and Maskin's (1989) weak renegotiation-proofness criterion) continue to admit multiple equilibria. Second, I have used hyperbolic discounting to develop a model of precommitment (Laibson (1993a)). Third, I am using hyperbolic discounting to develop a model of self-reward and mental accounts (Laibson (1993b)).

[14]The exception to this observation is the Phelps-Pollak equilibrium.

## Sources

Ainslie, George W. (1975) "Specious Reward: A Behavioral Theory of Impulsiveness and Impulsive Control." *Psychological Bulletin*, 82, 463-96.

Ainslie, George W. (1986) "Beyond Microeconomics. Conflict Among Interests in a Multiple Self as a Determinant of Value." in Jan Elster, ed., *The Multiple Self*, Cambridge: Cambridge University Press, 133-175.

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

Akerlof, George A. (1991) "Procrastination and Obedience." *American Economic Review*, AEA Papers and Proceedings, 1-19.

DeVilliers, P., and Richard J. Herrnstein (1976) "Toward a Law of Response Strength." *Psychological Bulletin*, 83, 1131-53.

Farrell, Joseph and Eric Maskin (1989) "Renegotiation in Repeated Games." *Games and Economic Behavior*. 1, 327-360.

Goldman, Steven M. (1980) "Consistent Plans." *Review of Economic Studies*, 47, 533-37.

Herrnstein, Richard J. (1961) "Relative and Absolute Strengths of Response as a Function of Frequency of Reinforcement." *Journal of the Experimental Analysis of Animal Behavior*, 4, 267-72.

Laibson, David I. (1993a) "Golden Eggs and Hyperbolic Discounting." MIT mimeo.

Laibson, David I. (1993b) "Mental Accounts, Self-Control, and an Intrapersonal Principal-Agent Problem." MIT mimeo.

Loewenstein, George, and Drazen Prelec. (1992) "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *Quarterly Journal of Economics*, 57:2, 573-598.

Peleg, Bezalel, and Menahem E. Yaari. (1973) "On the Existence of a Consistent Course of Action when Tastes are Changing," *Review of Economic Studies*, 40, 391-401.

Phelps, E. S., and R. A. Pollak. (1968) "On Second-best National Saving and Game-equilibrium Growth." *Review of Economic Studies*, 35, 185-199.

Pollak, R. A. (1968) "Consistent Planning." *Review of Economic Studies*, 35, 201-208.

Prelec, Drazen. (1989) "Decreasing Impatience: Definition and Consequences." Harvard Business School Working Paper.

Ramsey, Frank P. (1928) "A Mathematical Theory of Saving." *Economic Journal*, 38, 543-559.

Schelling, Thomas C. (1983) "Ethics, Law, and the Exercise of Self-Command." in Sterling M. McMurrin, ed., *The Tanner Lectures on Human Values IV*. Salt Lake City: University of Utah Press, 43-79.

Stokey, Nancy L., Robert E. Lucas Jr., and Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*. Cambridge MA: Harvard University Press.

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