POVERTY AND SELF-CONTROL

B. Douglas Bernheim
Stanford University and NBER

Debraj Ray
New York University

Sevin Yeltekin
Carnegie Mellon University

April 2012

NOTE: This version is missing a proof of the main result, and is extremely preliminary. Comments and suggestions welcome.

"When you ain't got nothin', you got nothin' to lose." *Bob Dylan*

1. Introduction

The absence of self-control is often cited as an important contributory cause of persistent poverty, particularly (but not exclusively) in developing countries. Recent research indicates that the poor not only borrow at high rates,[1] but also forego profitable small investments.[2] To be sure, traditional theory — based on high rates of discount and minimum subsistence needs — can take us part of the way to an explanation. But it cannot provide a full explanation, for the simple reason that

[1]Informal interest rates in developing countries are notoriously high; see, for example Aleem (1990). But even formal interest rates are extremely high; for instance, the rates charged by microfinance organizations. Bangladesh recently capped formal microfinance interest rates at 27% per annum, a restriction frowned upon by the *Economist* (2010). Banerjee and Mullainathan (2010) cite other literature and argue that such loans are taken routinely and not on an emergency basis.

[2]Goldstein and Udrey (1999) and Udry and Anagol (2006) document high returns to agricultural investment in Ghana, even on small plots, while Duflo, Kremer, and Robinson (2010) identify high rates of return to small amounts of fertliizer use in Kenya, and de Mel, McKenzie, and Woodruff (2008) demonstrate high returns to microenterprise in Sri Lanka. Banerjee and Duflo (2005) cite other studies that also show high rates of return to small investments.

the poor exhibit a documented desire for commitment.[3] The fact that individuals are often willing to pay for commitment devices such as illiquid deposit accounts or ROSCA participation suggests that time inconsistency and imperfect self-control are important explanations for low saving and high borrowing, complementary to those based on impatience, minimum subsistence or a failure of aspirations.

A growing literature already recognizes that the (in)ability to exercise self-control is central to the study of intertemporal behavior.[4] Our interest lies in how self-control and economic circumstances *interact*. If self-control (or the lack thereof) is a fixed trait, independent of personal economic circumstances, then the outlook for policy interventions that encourage the poor to invest in their futures – particularly one-time or short-term interventions – is not good. But another possibility merits consideration: poverty *per se* may damage self-control. If that hypothesis proves correct, then the chain of causality is circular, and poverty is itself responsible for the low self-control that perpetuates poverty.[5] In that case, policies that help the poor *begin* to accumulate assets may be highly effective, even if they are temporary.

The preceding discussion motivates the central question of this paper: is there some *a priori* reason to expect poverty to perpetuate itself by undermining an individual's ability to exercise self-control? Our objective requires us to define self-control formally and precisely. The term itself implies an internal mechanism, so we seek a definition that does not reference any externally-enforced commitment devices. Following Strotz (1956), Phelps and Pollak (1968) and others, we adopt the view that self-control problems arise from time-inconsistent preferences: the

[3]See, for example, Shipton (1992) on the use of lockboxes in Gambia, Benartzi and Thaler (2004) on employee commitments to save out of future wage increases in the United States, Ashraf, Karlan, and Yin (2006) on the use of a commitment savings product in the Philippines, Aliber (2001), Gugerty (2001, 2007) and Anderson and Baland (2002) on ROSCA participation as a commitment device, and Duflo, Kremer, and Robinson (2010) on fertilizer use in Kenya (specifically, adoption just after harvest). See also the theoretical contributions of Ambec and Treich (2007) and Basu (2010).

[4]See, for instance, Akerlof (1991), Ainslie (1992), Thaler (1992), Laibson (1997), or O'Donoghue and Rabin (1999). There are social aspects to the problem as well. Excess spending may be generated by discordance within the household (e.g., husband and wife have different discount factors) or by demands from the wider community (e.g., sharing among kin or community).

[5]Arguments based on aspiration failures generate parallel traps: poverty can be responsible for frustrated aspirations, which stifle the incentive to invest. See, e.g., Appadurai (2004), Ray (2006), Genicot and Ray (2009) and the recent *United Nations Development Program* Report for Latin America based on this methodology. However, this complementary approach does not generate a demand for commitment devices.

absence of self-control is on display when an individual is unable to follow through on a desired plan of action. What then constitutes the *exercise* of self-control? We take guidance from the seminal work of the psychologist George Ainslee (1975, 1992), who argued that people maintain personal discipline by adopting private rules (e.g., "never eat dessert"), and then construing local deviations from a rule as having global significance (e.g., "if I eat dessert today, then I will probably eat dessert in the future as well"). Formally, with time-inconsistent preferences, it is natural to model behavior as a subgame-perfect Nash equilibrium of a dynamic game played by successive incarnations of the single decision-maker.[6] For such a game, any equilibrium path is naturally interpreted as a personal rule, in that it describes the way in which the individual is supposed to behave. Moreover, history-dependent equilibria can capture Ainslee's notion that local deviations from a personal rule can have global consequences.[7] For example, in an intrapersonal equilibrium, an individual might correctly anticipate that violating the dictum to "never eat dessert" will trigger an undesirable behavioral pattern. Under that interpretation, the scope for exercising self-control is sharply defined by the set of outcomes that can arise in subgame-perfect Nash equilibria.

We assume that time-inconsistency arises from *quasi-hyperbolic discounting* (also known as $\beta\delta$-discounting), a standard model of intertemporal preferences popularized by Laibson (1994, 1996, 1997) and O'Donoghue and Rabin (1999). To determine the full scope for deliberate self-control, we study the features of the set of *all* subgame-perfect Nash equilibria. To avoid excluding any viable personal rules, we impose no restrictions whatsoever on strategies (we do not require stationarity, for instance, or that the decision-maker punish deviations by reverting to the Markov-perfect equilibrium). This approach contrasts with the vast majority of the existing literature, which focuses almost exclusively on Markov-perfect equilibria (which allow only for payoff-relevant state-dependence), thereby ruling out virtually all interesting personal rules.[8] By studying the entire class of subgame-perfect Nash equilibria, we can determine when an individual can exercise sufficient

[6]This approach is originally due to Strotz (1956).
[7]This interpretation of self-control has been offered previously by Laibson (1997), Bernheim, Ray, and Yeltekin (1999), and Benhabib and Bisin (2001). See Bénabou and Tirole (2004) for a somewhat different interpretation of Ainslee's theory.
[8]Exceptions include Laibson (1994), Bernheim, Ray, and Yeltekin (1999), and Benhabib and Bisin (2001).

self-control to accumulate greater wealth, and when she cannot. In particular, we can ask whether self-control is more difficult when initial assets are low, compared to when they are high.

The model we use is standard. There is a single asset which can be accumulated or run down at some fixed rate of return. By using suitably defined present values, all flow incomes are nested into the asset itself. The core restriction is a strictly positive lower bound on assets, to be interpreted as a credit constraint. In other words, the individual cannot instantly consume *all* future income. The lower bound may therefore be interpreted as referring to that fraction of present-value income which she cannot currently consume.

Apart from this lower bound, the model is constructed to be scale-neutral. We assume that individual payoff functions are homothetic, so we deliberately eliminate any preconceived relationship between assets and savings that is dependent on preferences alone. (We return to this point when connecting our model to related literature.) Discounting is quasi-hyperbolic.

It is notoriously difficult to characterize the set of subgame-perfect Nash equilibria (or equilibrium values) for all but the simplest dynamic games, and the problem of self-control we study here is, alas, no exception. We therefore initially examined our central question by solving the model numerically using standard tools. Figure 1 illustrates the results of one such exercise (which we explain at greater length later in the paper). The horizontal axis measures assets in the current period, and the vertical axis measures assets in the next period. Thus, points above, on, and below the 45 degree line indicate asset accumulation, maintenance, and decumulation respectively. In this exercise, there is an asset threshold below which *all* equilibria lead to decumulation. Thus, with low assets, it is impossible to accumulate assets by exercising self-control through any viable personal rule; on the contrary, assets necessarily decline until the individual's liquidity constraint binds. In short, we have a poverty trap.

However, above that threshold, there are indeed viable personal rules that allow the individual to accumulate greater assets. Moreover, as we will see later, the most attractive equilibria starting from above the critical threshold lead to unbounded accumulation.

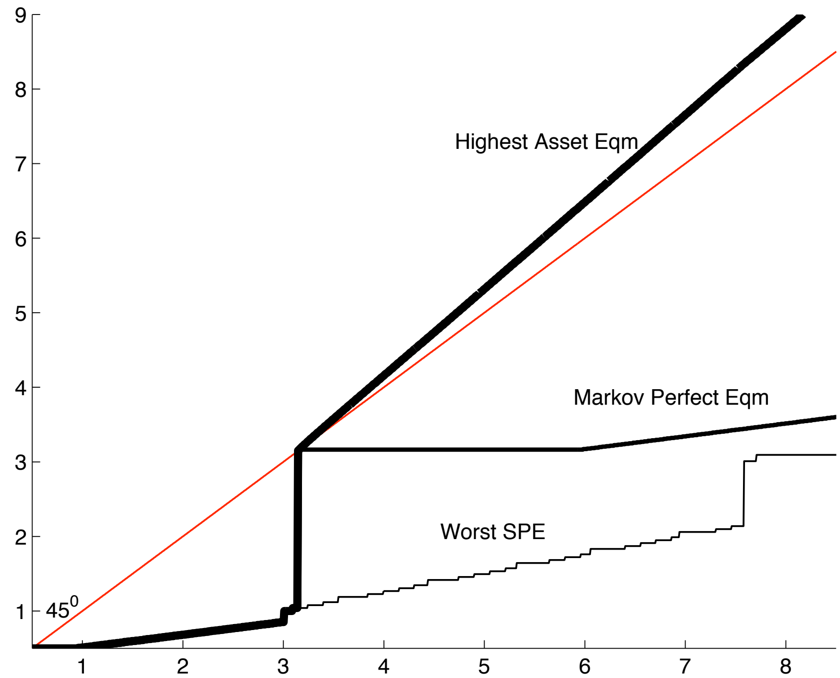


FIGURE 1. ACCUMULATION AT DIFFERENT ASSET LEVELS.

The example suggests both our central conjecture and a (deceptively) simple intuitive argument in support of it. If imperfect capital markets impose limits on the extent to which an individual can borrow against future income, then potential intrapersonal "punishments" (that is, the consequences of deviating from a personal rule) cannot be all that bad when assets are low. If these mild repercussions are suitably anticipated, an individual will fail to exercise self-control. However, when an individual has substantial assets, she also has more to lose from undisciplined future behavior, and hence potential punishments are considerably more severe (in relative terms). So an individual would be better able to accumulate additional assets through the exercise of self-control when initial assets are higher. Obviously, if time inconsistency is sufficiently severe, decumulation will be unavoidable regardless of initial assets, and if it is sufficiently mild, accumulation will be possible regardless of initial assets (provided the individual is sufficiently patient). But for intermediate degrees of time inconsistency, we would expect decumulation to be unavoidable with low assets, and accumulation to be feasible with high assets.

It turns out, however, that the problem is considerably more complicated than this simple intuition suggests. (The overwhelmingly numerical nature of our earlier working paper, Bernheim, Ray, and Yeltekin (1999), bears witness to this assertion.) The credit constraint at low asset levels infects individual behavior at *all* asset levels. In particular, they affect the structure of "worst personal punishments" in complex ways, as assets are scaled up. The example of Figure 1 illustrates this point quite dramatically: there are asset levels at which the *lowest* level of continuation assets jumps up discontinuously. As assets cross those thresholds, the worst punishment becomes *less* rather than *more* severe, contrary to the intuition given above. Thus, on further reflection, it is not at all clear that the patterns exhibited in Figure 1 will emerge more generally.

Our main theoretical result demonstrates, nevertheless, that the central qualitative properties of Figure 1 are quite general. For intermediate degrees of time inconsistency (such that accumulation is feasible from some but not all asset levels), there is a threshold asset level below which accumulation is impossible, and above which decumulation is avoidable. There is always an asset level below which liquid wealth is exhausted in finite time (and hence a poverty trap), as well as a level above which the most attractive equilibria give rise to unbounded accumulation.

One might object to our practice of examining the entire set of subgame-perfect equilibria on the grounds that many such equilibria may be unreasonably complex. On the contrary, we show that worst punishments have a surprisingly simple "stick-and-carrot" structure:[9] following any deviation from a personal rule, the individual consumes aggressively for one period, and then returns to an equilibrium path that maximizes her (equilibrium) payoff *exclusive* of the hyperbolic factor. Thus, all viable personal rules can be sustained without resorting to complex forms of history-dependence.[10]

[9]Though there is an obvious resemblance to the stick-and-carrot punishments in Abreu (1988), the formal structure of the models and the arguments differ considerably. Most obviously, Abreu considered simultaneous-move repeated games, rather than sequential-move dynamic games with state variables.

[10]Indeed, Markov equilibria in this model appear to be more "complex", despite their "simple" dependence on just the payoff-relevant state. They typically involve several jump discontinuities, and their payoff behavior as a function of initial assets, suitably normalized, is often nonmonotonic. Also, identifying Markov equilibria is more computationally challenging than determining the key features of subgame-perfect equilibria.

Our analysis has a number of provocative implications for economic behavior and public policy. First (and most obviously), the relationship between assets and self-control argues for the use of "pump-priming" interventions that encourage the poor to start saving, and rely on self-control to sustain frugality at higher levels of assets. Second, policies that improve access to credit (thereby relaxing liquidity constraints) reduce the level of assets at which asset accumulation becomes feasible, thereby helping more individuals to become savers (although those who fail to make the transition fall further into debt). Intuitively, with greater access to credit, the consequences of a break in discipline become more severe, and hence that discipline is easier to sustain to begin with. Third, the opportunities to make commitments may be significantly less valuable to those with self-control problems than previous analyses have implied. For example, in certain circumstances, individuals with self-control problems will avoid opportunities to lock up funds (e.g., in retirement accounts or fixed deposit schemes), even when they wish to save. This occurs when desired saving exceeds the maximum amount that can be locked up, but not by too much. In such cases, locking up funds moderates the consequences of a lapse in discipline, thereby making self-control more difficult to sustain. Finally, we argue that the model can potentially provide an explanation for the observed "excess sensitivity" of consumption to income.

As noted above, we build on an earlier unpublished working paper (Bernheim, Ray, and Yeltekin (1999)), which made our main points through simulations, but did not contain theoretical results. Our work is most closely related to that of Banerjee and Mullainathan (2010), who also argue that self-control problems give rise to low asset traps. Though the object of their investigation is similar, their analysis has little in common with ours. They examine a novel model of time-inconsistent preferences, in which rates of discount differ from one good to another. "Temptation goods" (those to which greater discount rates are applied) are inferior by assumption; this assumed non-hometheticity of preferences automatically builds in a tendency to dissave when resources are limited, and to save when resources are high.

It is certainly of interest to study poverty traps by hardwiring non-homothetic self-control problems into the structure of preferences. Whether a poor person spends proportionately more on temptation goods than a rich person (alcohol versus IPods, say) then becomes an empirical matter. But we avoid such hardwiring entirely by studying homothetic preferences in an established model of time-inconsistency. The phenomena we study are traceable to a single built-in asymmetry: an imperfect credit market. Every scale effect in our setting arises from the interplay between credit constraints and the incentive compatibility constraints for personal rules. The resulting structure, in our view, is compelling in that it requires no assumption concerning preferences that must obviously await further empirical confirmation. In summary, though both theories of poverty traps invoke self-control problems, they are essentially orthogonal (and hence potentially complementary): Banerjee and Mullainathan's analysis is driven by assumed scaling effects in rewards, while ours is driven by scaling effects in punishments arising from assumed credit market imperfections.[11]

The rest of the paper is organized as follows. Section 2 presents the model, and Section 3 characterizes the set equilibrium values. Section 4 defines self control, and Section 5 studies the relationship between self-control and the initial level of wealth. Section 6 describes some implications of the theory. Section 7 conducts numerical exercises to supplement the analytical findings. Section 8 presents conclusions and some directions for future research. Proofs are collected in an Appendix.

2. Model

2.1. **Feasible Set and Preferences.** The feasible set is given by a restriction that links current assets, current consumption and future assets:

$$c_t = A_t - (A_{t+1}/\alpha),$$

as well as a lower bound on assets

$$A_t \geq B > 0,$$

and an initial asset level $A_0 = A \geq B$. Our leading interpretation of the lower bound is that it is a credit constraint. For instance, if F_t stands for *financial* assets at date t and y for income at every date, then A_t is the present value of financial

[11]While our model is also related to Laibson (1994) and Benhabib and Bisin (2001), our analysis is not. These two papers consider history-dependent strategies in a *fully* scalable model, in which both preferences are homothetic and there is no credit constraint. It follows, as we observe more formally below, that every equilibrium path can be replicated, by scaling, at all levels of initial assets, so that there is no relationship between poverty and self-control.

and labor assets:

$$A_t = F_t + \frac{\alpha y}{\alpha - 1}.$$

If credit markets are perfect, the individual will have all of A_t at hand today, and $B = 0$. We are not directly interested in this case (our analysis presumes $B > 0$) but it is easy enough to do so; see Laibson (1994). On the other hand, if she can borrow only some fraction $(1 - \lambda)$ of lifetime income, then $B = \lambda\alpha y/(\alpha - 1)$.[12]

An individual is equipped with quasi-hyperbolic preferences: her lifetime utility is given by

$$u(c_0) + \beta \sum_{t=1}^{\infty} \delta^t u(c_t),$$

where $\beta \in (0, 1)$ and $\delta \in (0, 1)$. We assume throughout that u has the constant-elasticity form

$$u(c) = \frac{c^{1-\sigma}}{1 - \sigma}$$

for $\sigma > 0$, with the understanding that $\sigma = 1$ refers to the logarithmic case $u(c) = \ln c$.

There is a good reason for the use of the constant-elasticity formulation. We wish our problem to be entirely scale-neutral in the absence of the credit constraint, so as to isolate fully the effect of that constraint. While we don't formally analyze the case in which $B = 0$, it is obvious that scale-neutrality is achieved there: any path with perfect credit markets can be freely scaled up or down with no disturbance to its equilibrium properties. In short, every scale effect in this model will arise from the interplay between credit constraints and the incentive compatibility constraints for personal rules.

2.2. **Restrictions on the Model.** The *Ramsey program* from A is the feasible asset stream that maximizes

$$\sum_{t=0}^{\infty} \delta^t \frac{c_t^{1-\sigma}}{1 - \sigma},$$

with initial stock $A_0 = A$. It is constructed without reference to the hyperbolic factor β.

[12]Another interpretation of B is that it is an investment in some fixed illiquid asset. We return to this interpretation when we talk about policy implications.

The Ramsey program is well-defined provided that utilities do not diverge, for which we assume throughout that

$$\mu \equiv \delta^{1/\sigma}\alpha^{(1-\sigma)/\sigma} < 1. \tag{1}$$

We will also be interested in situations in which the Ramsey optimum exhibits growth, which imposes

$$\delta\alpha > 1. \tag{2}$$

Under (1) and (2), it is easy to see that along the Ramsey program,

$$c_t = (1-\mu)A_t,$$

assets grow exponentially:

$$A_{t+1} = A_t\left(\delta^{1/\sigma}\alpha^{1/\sigma}\right)^t = A_t\left(\mu\alpha\right)^t,$$

and the value of the Ramsey program from A — call it $R(A)$ — is finite.

Before we turn to a precise definition of equilibrium, note that when $\sigma \geq 1$, utility is unbounded below and it is possible to sustain all sorts of outcomes by taking recourse to punishments that either impose zero consumption or a progressively more punitive sequence of non-zero consumptions (see Laibson (1994) for a discussion of this point). These punishments rely on the imposition of unboundedly negative utility. We find such actions unrealistic, and eliminate them by assuming that consumption is bounded below at every asset level. More precisely, we assume that at every date,

$$c_t \geq \nu A_t, \tag{3}$$

where ν is to be thought of as a small but positive number. It is formally enough to presume that $\nu < 1-\mu$, so that Ramsey accumulation can occur unhindered, but the reader is free to think of this bound as tiny.

2.3. **Equilibrium.** A *history* h_t at date t is any vector of assets $(A_0, \ldots, A_t)$ with $A+0 \geq B$ and $B \leq A_{s+1} \leq \alpha A_s$ for every $s = 0, \ldots, t-1$. A *policy* ϕ specifies a feasible continuation asset $\phi(h_t)$ following every history. For any history h_t, let $A(h_t)$ be the last entry in h_t; i.e., the initial asset at the start of date t. If h_t is a history and x is a feasible asset choice, i.e., $B \leq x \leq \alpha A(h_t)$, let $h_t.x$ denote the corresponding "continuation history". A policy ϕ defines a *value* V_ϕ by

$$V_\phi(h_t) \equiv \sum_{s=t}^{\infty} \delta^{s-t} u\left(A(h_s) - \frac{\phi(h_s)}{\alpha}\right),$$

where h_s (for $s > t$) is recursively defined from h_t by $h_{s+1} = h_s.\phi(h_s)$ for $s \geq t$. Similarly, ϕ also defines a *payoff* P by

$$P_\phi(h_t) \equiv u\left(A(h_t) - \frac{\phi(h_t)}{\alpha}\right) + \beta V_\phi(h_t.\phi(h_t)),$$

for every history h_t. Note that values exclude the hyperbolic factor β, while payoffs include them.

An *equilibrium* is a policy such that at every history h_t,

$$P_\phi(h_t) \geq u\left(A(h_t) - \frac{x}{\alpha}\right) + \delta\beta V_\phi(h_t.x), \tag{4}$$

for every $x \in [B, \alpha A(h_t)]$.

We augment policies by allowing for public randomization at any asset level. That makes it possible for agents to convexify the set of continuation asset values by randomize between any two such values. So the set of equilibrium (continuation) values at any asset level A is an interval.

3. Existence and Characterization of Equilibrium

For each initial asset level $A \geq B$, let $\mathcal{V}(A)$ be the set of all equilibrium values available at asset level A. If $\mathcal{V}(A)$ is nonempty, let $H(A)$ and $L(A)$ be its supremum and infimum values. It is obvious from our assumed lower bound on consumption and from utility convergence (see (1)) that

$$-\infty < L(A) \leq H(A) \leq R(A) < \infty,$$

where $R(A)$ is the Ramsey value defined earlier. Indeed, once we rule out unrealistic cascades of punishments that generate arbitrarily negative utility, a tighter and more intuitive bound is available for worst punishments:

Observation **1.** *Suppose that $\mathcal{V}(A)$ is nonempty for every $A \geq B$. Then*

$$L(A) \geq u\left(A - \frac{B}{\alpha}\right) + \frac{\delta}{1-\delta} u\left(\frac{\alpha - 1}{\alpha} B\right) \tag{5}$$

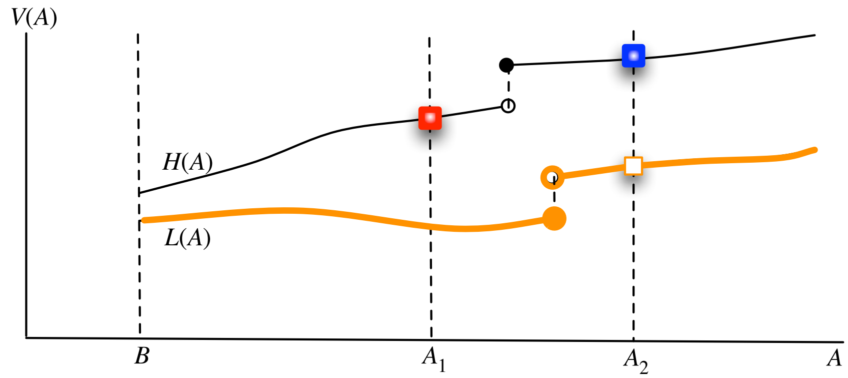


FIGURE 2. EQUILIBRIUM VALUES.

Notice how Observation 1 kicks in as long as we place *any* (small) lower bound on consumption, as described in (3).[13] It gives us an anchor to iterate a self-generation map, both for analytical use and for equilibrium computation. To this end, consider a nonempty-valued correspondence $\mathcal{W}$ on $[B, \infty)$ such that for all $A \geq B$,

$$\mathcal{W}(A) \subseteq \left[u\left(A - \frac{B}{\alpha}\right) + \frac{\delta}{1-\delta} u\left(\frac{\alpha - 1}{\alpha} B\right), R(A) \right]. \tag{6}$$

Say that $\mathcal{W}$ *generates* the correspondence $\mathcal{W}'$ if for every $A \geq B$, $\mathcal{W}'(A)$ is the collection of *all* values W such there is a feasible asset choice x and $V \in \mathcal{W}(x)$ with

$$W = u\left(A - \frac{x}{\alpha}\right) + \delta V, \tag{7}$$

while for every feasible x', there is $V' \in \mathcal{W}(x')$ such that

$$u\left(A - \frac{x}{\alpha}\right) + \beta\delta V \geq u\left(A - \frac{x'}{\alpha}\right) + \beta\delta V'. \tag{8}$$

Given Observation 1 and the Ramsey upper bound on equilibrium values, standard arguments tell us that the equilibrium correspondence $\mathcal{V}$ generates itself, and indeed, it contains any other correspondence that does so.

[13] In contrast, if there is no lower bound, and utilities are unbounded below, we can generate artificially low punishments by effectively using "Ponzi threats".

Define a sequence of correspondences on $[B, \infty)$, $\{\mathcal{V}_k\}$, by

$$\mathcal{V}_0(A) = \left[u\left(A - \frac{B}{\alpha}\right) + \frac{\delta}{1-\delta} u\left(\frac{\alpha - 1}{\alpha} B\right), R(A) \right].$$

for every $A \geq B$, and recursively, $\mathcal{V}_k$ generates $\mathcal{V}_{k+1}$ for all $k \geq 0$. It is obvious that the graph of $\mathcal{V}_k$ contains the graph of $\mathcal{V}_{k+1}$. We assert

Proposition 1. *An equilibrium exists from any initial asset level, so that the equilibrium correspondence $\mathcal{V}$ is nonempty-valued. Moreover, for every $A \geq B$,*

$$\mathcal{V}(A) = \bigcap_{k=0}^{\infty} \mathcal{V}_k(A). \tag{9}$$

and $\mathcal{V}$ has closed graph.

This proposition is not standard, or at least it does not follow from any known general arguments, for two reasons. First, it (quickly) proves existence of equilibrium, though the method used may not apply more generally to all games with state variables.[14] Second, the closed-graph property does not follow from the nested compact sets argument, because the sets in question (the graphs of $\mathcal{V}_k$) are not compact.

Figure 2 illustrates a set of equilibrium values. Imagine supporting the highest possible value $H(A_1)$ at asset level A_1. That might require the choice of asset A_2 — that would be the stipulation of the equilibrium policy — followed by the continuation value $H(A_2)$. Any other choice would be followed by other continuation values designed to discourage that choice, so that the inequality in (4) holds. The figure illustrates the "best" way of doing this under the presumption that the equilibrium value set is compact-valued and has closed graph: simply choose the worst continuation value $L(x)$ if $x \neq A_1$.

4. Self Control

Viewed in the spirit of Ainslee's definition, self-control refers to a *possibility*; that is, to a feature of *some* element of the equilibrium correspondence. One might ask, for instance, if the unfettered Ramsey outcome is possible at some equilibrium. But that requires that the agent entirely transcend her hyperbolic urges. All other

[14] For a longer but more general argument, see Goldman (1980) and Harris (1985).

attempts, including indefinite accumulation at rates very close to the first-best outcome, must then be deemed a failure (of self-control), which we find too strong. We therefore begin with a weak definition: say there is *self-control* at asset level A if the agent is capable of positive saving at A in *some* equilibrium.

To be sure, we might be interested in whether the individual is capable of indefinite accumulation. Say that there is *strong self-control* at asset level A if the agent is capable of unbounded accumulation — i.e., $A_t \to \infty$ — for some equilibrium path emanating from A.

Now we look at the flip side of self-control. Clearly, we must define the absence of self-control as a situation in which accumulation isn't possible under *any* equilibrium. But that failure is compatible with several outcomes: the stationarity of assets, a downward spiral of assets to a lower level that nevertheless exceeds the lower bound, or a progressive downward slide all the way to the minimal level B. In a symmetric way, we single out two features: say that *self control fails* at A if every equilibrium continuation asset is strictly smaller than A, and more forcefully, that there is a *poverty trap at asset* A if in every equilibrium, assets decline over time from A to the lower bound B.

There is intermediate ground between strong self-control and a poverty trap: it is, in principle, possible for an agent to be incapable of indefinite accumulation, while at the same time she can avoid the poverty trap.

That said, there are situations in which self-control is possible at *all* asset levels. For instance, if β is close to 1, there is (almost) no time-inconsistency and all equilibria should exhibit accumulation, given our assumption that the Ramsey program involves indefinite growth. Conversely, if the agent exhibits a high degree of hyperbolicity (β small), there may be a failure of self-control no matter what asset level we consider. Call a case *uniform* if there are no switches: either there is no failure of self control at every asset level, or there is no self-control at every asset level.

A good example of uniformity is given by the case in which credit markets are perfect. While we don't formally study perfect credit markets in this paper, the assertion is obvious: if continuation asset x can be sustained at asset level A, then continuation asset λx can be just as easily sustained when the asset level is λA, for any $\lambda > 0$. Indeed, we've deliberately constructed the model in this fashion, so to understand better the "direction of scale bias" created by introducing imperfect credit markets.

In particular, the *nonuniform* cases are of interest to us. These are situations in which self-control is possible at some asset level A, while there is a failure of self-control at some other asset level A'. Whether A' is larger or smaller than A, or indeed, whether there could be several switches back and forth, are among the central issues that we wish to explore. It should be added that while we do not have a full characterization of when a case is nonuniform, such cases exist in abundance (and this can be confirmed by numerical analysis).

We close this section with an intuitive yet nontrivial characterization of self control. Because the equilibrium value correspondence has closed graph, the *largest continuation asset* is well-defined: it is the largest value of equilibrium asset $X(A)$ sustainable at A. The closed-graph property of Proposition 1 guarantees that $X(A)$ is well-defined and usc, and a familiar single-crossing argument tells us that it is nondecreasing. Note that $X(A)$ isn't necessarily the value-maximizing choice of asset; it could well be higher than that. Yet $X(A)$ is akin to a sufficient statistic that can be used to characterize all the concepts in this section.

PROPOSITION **2.** (i) *Self control is possible at A if and only if $X(A) > A$.*

(ii) *Strong self control is possible at A if and only if $X(A') > A'$ for all $A' \geq A$.*

(iii) *There is a poverty trap at A if and only if $X(A') < A'$ for all $A' \in (B, A]$.*

(iv) *There is uniformity if and only if $X(A) \geq A$ for all $A \geq B$, or $X(A) \leq A$ for all $A \geq B$.*

Parts (i) and (iv) of the proposition are obvious, but parts (ii) and (iii), while intuitive, are not. For instance, part (ii) will need two steps; see the formal proof: (a) that the X's — or some path that accumulates — can be recursively "strung together", and (b) that such a path does not converge to some bounded limit on the 45^0 line (after all, there is no guarantee that X is continuous[15]). That said, the proposition is true and it will help us in visualizing the proof of the main theorem.

[15] That is, the assertion $X(A') > A'$ for all $A' \geq A$ is fully compatible with the possibility that $X(A)$ converges to A from below as $A \uparrow \bar{A}$ for some $\bar{A} < \infty$, while at the same time $X(\bar{A}) > \bar{A}$.

5. Initial Assets and Self-Control

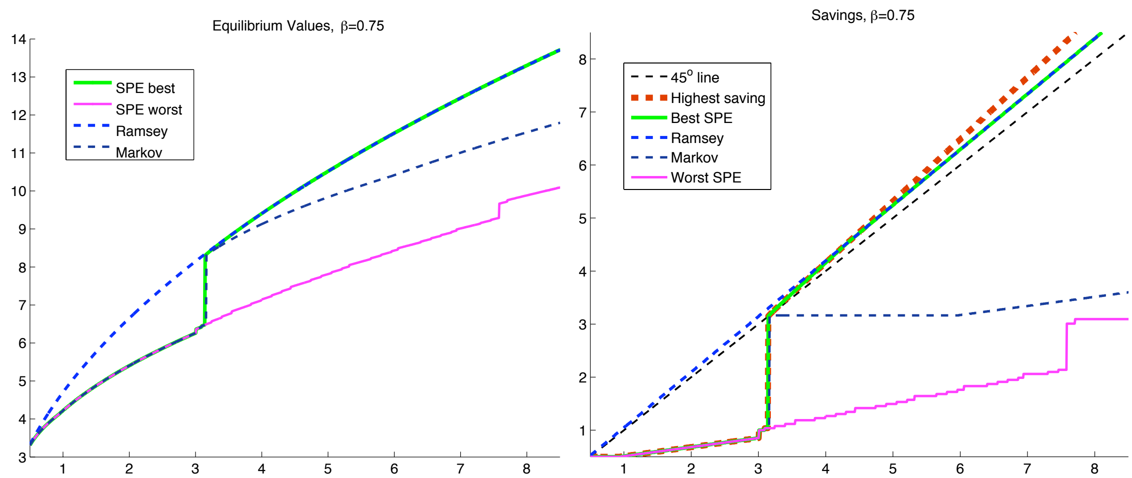
It is obvious that if $B > 0$, then "scale-neutrality" fails. For instance, at asset level B, it isn't possible to decumulate assets (by assumption), while that may be an equilibrium outcome at $A > B$. This trivial failure of neutrality opens the door to all sorts of more interesting failures of neutrality. For instance, accumulation at some asset level A may be sustained by the threat of decumulation in the event of non-compliance; such threats will not be credible at asset levels close to B.

Such internal checks and balances are not merely technical, but descriptive (we feel) of individual ways of coping with commitment problems. One coping mechanism is, of course, "external": an individual might commit to a fixed deposit account if available, or even accounts that force her to make regular savings deposits in addition to imposing restrictions on withdrawal. We will have more to say about such mechanisms below. But the other coping mechanism is "internal": an agent might react to an impetuous expenditure on her part by going on a continued though possibly temporary binge. We impose the condition that such a binge must be a valid continuation equilibrium. If so, the threat of a credible binge might help to keep the agent in check.

With this "internal mechanism" in mind, let's ask the question of how high assets might help an individual to exhibit self-control. The success of such an enterprise must depend on the severity of the consequences following an impetuous act of current consumption. One simple intuition is that those consequences are more severe when the individual has more assets, and hence more to lose. But we know that such an argument can run either way, for several reasons.[16]

A specific problem that arises here is that the "severity of punishment" (even if suitably normalized) isn't monotonic in assets. Figure 3 contains a numerical example which makes the point easily. The left panel shows various value selections from the equilibrium correspondence, as also the Ramsey value. The lowest selection is $L(A)$. It jumps several times; the diagram shows one such jump between

[16]For instance, in moral hazard problems with limited liability, a poor agent might face more serious incentive problems than a rich one; see, e.g., Mookherjee (1997). On the other hand, the curvature of the utility function will permit the inflicting of higher *utility* losses on poorer individuals, alleviating moral hazard and conceivably permitting the poor to be better managers (Banerjee and Newman (1991)).



(A) Values (B) Policies

FIGURE 3. JUMPS IN VALUES AND POLICIES

the values 7 and 8. The right hand panel describes the corresponding choice of continuation assets. The jump in $L(A)$ shows that in general, punishment values (even after normalizing by the higher asset values) cannot be decreasing in A.

The jump in L is related to the failure of lower hemicontinuity of the constraint set in the implicit minimization problem that defines lowest values. That constraint set is constructed from the graph of the equilibrium value correspondence, in which all continuation values must lie. As assets converge down to some limit, discontinuously lower values may become available, and as the numerical example illustrates, this phenomenon cannot be ruled out in general.

Nevertheless, worst punishments have a relatively simple structure, and this is what we turn to next.

5.1. **Worst Punishments.** We will show that worst punishments involve a single spell of binging, followed (roughly) by a return to the best possible continuation value. To formalize this notion, define, for any $A > B$, $H^{-}(A)$ by the left limit of $H(A^n)$ as A^n converges up to A, with $A^n < A$ for all n. This is a well-defined concept because H is obviously nondecreasing and therefore has limits from the left.

Proposition **3.** *Let A' be a choice of continuation asset and V' a choice of continuation value that minimizes equilibrium value at A_*; i.e.,*

$$L(A_*) = u\left(A_* - \frac{A'}{\alpha}\right) + \delta V'.$$

Then, if $A' > B$,

$$V' \geq H^-(A').$$

The proof of this proposition is simple and instructive enough to be included in the main text. We follow it up with a discussion.

Proof. Suppose that the assertion is false. Figure 4 illustrates this by sketching in the asset choice $A' > B$ and continuation V' that presumably minimizes equilibrium value; note that $V' < H^-(A')$ in the diagram in defiance of the proposition. Reduce A' slightly to A'' and note that $H(A'') > V'$, as it must be if $V' < H^-(A')$, as presumed. It follows that the resulting *payoff*

$$u\left(A_* - \frac{A''}{\alpha}\right) + \beta\delta H(A'')$$

exceeds the corresponding payoff when (A', V') are chosen. On the other hand, by the very fact that the payoff under (A', V') is an *equilibrium* payoff, it must be that

$$u\left(A_* - \frac{A''}{\alpha}\right) + \beta\delta L(A'')$$

is no higher than our payoff at (A', V'). Because all continuation values between $L(A'')$ and $H(A'')$ are available by public randomization, it follows that there exists an equilibrium value $V'' \in [L(A''), H(A'')]$ such that

$$u\left(A_* - \frac{A''}{\alpha}\right) + \beta\delta V'' = u\left(A_* - \frac{A'}{\alpha}\right) + \beta\delta V'. \tag{10}$$

Two conclusions follow from (10). First, the specification (A'', V'') must also be part of some equilibrium play from A_*. Second, the value of this play is

$$\begin{aligned} u\left(A_* - \frac{A''}{\alpha}\right) + \delta V'' &= u\left(A_* - \frac{A'}{\alpha}\right) + \delta V' + \delta(1-\beta)(V'' - V') \\ &< u\left(A_* - \frac{A'}{\alpha}\right) + \delta V', \end{aligned} \tag{11}$$

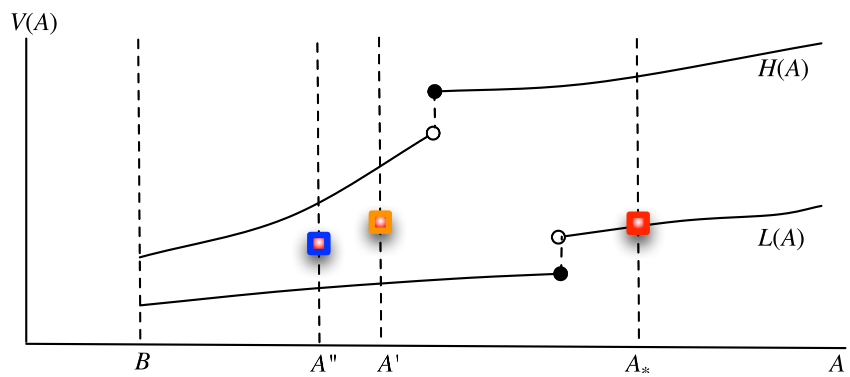


FIGURE 4. AN ILLUSTRATION TO PROVE PROPOSITION 3.

where the last inequality follows from $A'' < A'$ and (10), which together imply $V'' < V'$.

But (11) contradicts the supposition that the play (A', V') achieves the lowest equilibrium value at A_*. ■

The proposition states that worst punishments involve precisely one period of "excessive consumption", not counting the presumed binge that led to the deviation in the first place, followed by a reversion to the best possible equilibrium value — or approximately so[17] — from the asset level after the binge.

It is important to note that optimal punishments have an extremely simple structure: no complex rules are followed that might justify a retreat to "simpler" equilibrium notions, such as Markov. Second, the structure of the punishment path is plausible. An individual doesn't fall of the wagon forever, but there is still retribution for a deviation: a binge is followed by a further binge, the fear of which acts as a deterrent. After that, the individual is back on the wagon. Finally, there is a sense in which these punishments are renegotiation-proof: while the earlier, deviating self fears the low-value punishment path, the self that inflicts the punishment is actually treated rather well: he gets to enjoy a free binge, followed by the promise of self-control being exercised in the future.

[17]We should add that this is an approximate description. The reversion is to the best value provided that the asset level post-binge is strictly higher than B, and provided that the best value selection is continuous at that asset level. Otherwise the return is not necessarily to the best equilibrium continuation: recall the definition of H^-.

Our punishment profile is reminiscent of the carrot-and-stick property for optimal penal codes in repeated games (Abreu (1988)), but of course there is no reason why that property should hold, in general, for games with state variables, of which our model is an example. Indeed, the particular reason for the carrot-and-stick-like structure in our context comes from the hyperbolic factor β. That parameter dictates that the most effective punishments are achieved by as much excess consumption "as possible" in the very first period of the punishment. From the point of view of the deviator, that first period lies in *his* future, and as such it is a bad prospect (hence an effective punishment). From the point of view of the punisher, the punishment might actually yield pleasing equilibrium payoffs, given that a current binge followed by a high continuation value is something that quite appeals to him. That is, the carrot-and-stick feature is very much in the eyes of the deviator, and not in the eye of the punisher, a distinction that is often not present in many repeated games; see, e.g., Abreu (1986).

5.2. **The Relationship Between Wealth and Self-Control.** The possibility that worst equilibrium values can abruptly rise with wealth leads to the nihilistic suspicion that no general connection can be made between wealth and self-control. Indeed, this factor explains the absence of any analytical results in our 1999 paper. Nevertheless, not one of the extensive numerical examples that we have studied bears out this suspicion. Bernheim, Ray and Yeltekin (1999) show that either we are in one of the two uniform cases (equilibrium accumulation possible everywhere, or or equilibrium accumulation impossible anywhere), or the situation looks quite generically like Figure 3. Initially, there is asset decumulation in every equilibrium, followed by the crossing of a threshold at which indefinite accumulation becomes possible. The non-uniform cases invariably display a failure of self-control to begin with (at low asset levels), followed by the emergence and maintenance of self-control after a certain asset threshold has been crossed.

The main proposition of this paper supports the numerical analysis:

Proposition **4.** *In any non-uniform case:*

(i) *There is* $A_1 > B$ *such that every* $A \in [B, A_1)$ *exhibits a poverty trap.*

(ii) *There is* $A_2 \geq A_1$ *such that every* $A \geq A_2$ *exhibits strong self-control.*

The proposition states that in any situation where imperfect credit markets are sufficient to disrupt uniformity, the lack of scale neutrality manifests itself in a particular way. At low enough wealth levels, individuals are unable to exert self-control, and their actions must generate a poverty trap. At high enough wealth levels, indefinite accumulation is possible. There is, of course, no reason *a priori* why this must be the case. It is possible, for instance, that there a maximal asset level beyond which accumulation ceases altogether, or that there are (infinitely) repeated intervals along which accumulation and decumulation occur alternately. But the proposition rules out these possibilities.

This proposition provides partial vindication for the numerical analysis conducted in Bernheim, Ray and Yeltekin (1999). One should compare this finding with the main observation in Banerjee and Mullainathan (2010). They make the same point as we do here and (numerically) in our earlier work. Self-control problems give rise to low asset traps. But the analysis is different. They study time-inconsistent preferences over multiple goods in which rates of discount differ from one good to another. "Temptation goods" have higher discount rates attached to them. They are taken to be inferior by assumption. This assumed non-hometheticity of preferences generates a tendency to dissave when resources are limited. Our non-homotheticity manifests itself not via preferences but through the imperfection of credit markets. As we've discussed, there is no *a priori* presumption regarding the direction of that non-homotheticity.

In fact, the proposition fails to establish the existence of a *unique* asset threshold beyond which there is self-control, and below which there isn't. A demonstration of this stronger result is hindered in part by the possibility that worst punishments can move in unexpected ways with the value of initial assets. In fact, we conjecture that a "single crossing" is possibly not to be had, at least under the assumptions that we have made so far. From this perspective, the fact that "ultimately" all multiple crossings must cease — which is part of the assertion in the proposition — appears surprising, and the remainder of this section is devoted to an informal exposition of the proof.

5.3. **An Informal Exposition of the Main Proposition.** As we've mentioned on several occasions, it is the presence of imperfect credit markets that destroys scale-neutrality in our model. (The constant elasticity of preferences assures us

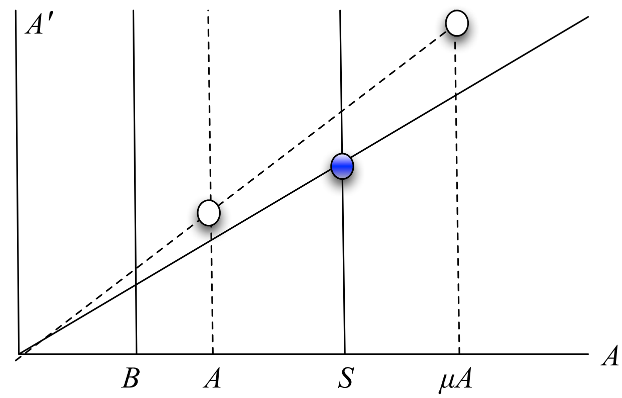


FIGURE 5. MODIFIED SCALING.

that otherwise, the situation would be fully scale-neutral.) Yet modified versions of scale-neutrality survive. One that is particularly germane to our argument is given in Lemma 1 below. To state it, define an asset level $S \geq B$ to be *sustainable* if there exists an equilibrium that permits indefinite maintenance of S.

It is important to appreciate that a sustainable asset level need not permit strict accumulation, and more subtly, an asset level that permits strict accumulation *need not* be sustainable.[18]

LEMMA **1.** *Let $S > B$ be a sustainable asset level. Define $\mu \equiv S/B > 1$. Then for any initial asset level $A \geq B$, if continuation asset A' can be supported as an equilibrium choice, so can the continuation asset $\mu A'$ starting from μA.*

Figure 5 illustrates the lemma. First think of S as a new lower bound on assets. Then it is plain that any equilibrium action under the old lower bound B can be simply scaled up using the ratio of S to B, which is μ. Indeed, if we replaced the word "sustainable" by the phrase "physical minimum", then the lemma would be trivial. We would simply scale up *all* continuation actions from the old equilibrium specification to the new one. However, S is not a physical minimum, and deviations to lower asset levels are available, but there is no version of such a deviation in the earlier equilibrium that can be rescaled (deviations below B are not allowed, after all). Nevertheless, the formal proof of Lemma 1 shows that given the concavity of the utility function, such deviations can be suitably deterred. Thus, while S isn't a

[18]The continuation values created by continued accumulation might incentivize accumulation from A, while a stationary path may not create enough incentives for self-sustenance.

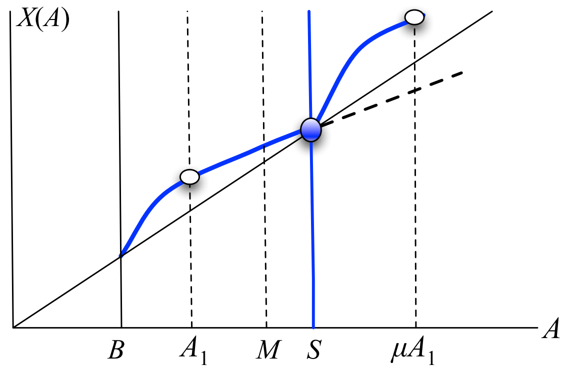


FIGURE 6. ESTABLISHING THE EXISTENCE OF A POVERTY TRAP.

physical lower bound, it permits us to carry out the same scaling we would achieve if it were.

Let's use Lemma 1 to establish the first part of the proposition:

(i) There is $A_1 > B$ such that every $A \in [B, A_1)$ exhibits a poverty trap.

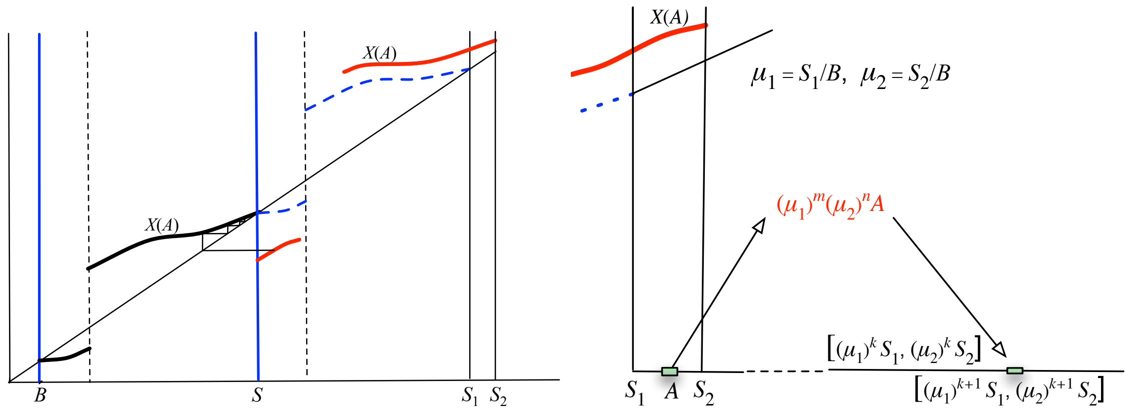
Recall that $X(A)$ to be the largest continuation asset in the class of all equilibrium outcomes at A. By Proposition 2, we will need to show that there is an asset level $A_1 > B$ such that $X(A) < A$ for all $A \in (B, A_1)$. Suppose, now, that the proposition is false; then — relegating the impossibility of eternal wiggles back and forth to the more formal arguments — there is $M > B$ such that $X(A) \geq A$ for all $A \in [B, M]$. Figure 6 illustrates this scenario.

Because we are in a non-uniform case, there exists an asset level A^* at which self-control fails, so by Proposition 2, $X(A^*) < A^*$. Let S be the supremum value of assets over $[B, A^*]$ for which $X(A) \geq A$. Note that *at* S, it must be the case that $X(S)$ *equals* S.[19] Because $X(S) = S$, S must be a sustainable asset, though this needs a formal argument.[20]

Now Lemma 1 kicks in to assert that $X(A)$ must exceed A just to the right of S: simply scale up the value $X(A_1)$ for some A_1 close to B to the corresponding value $\mu X(A_1)$ at μA_1, where $\mu \equiv S/B$. But that is a contradiction to the way we've

[19]It can't be strictly lower, for then X would be jumping down at S, and it can't be strictly higher for then we could find still higher asset levels for which $X(A) \geq A$.

[20]After all, it isn't *a priori* obvious that "stitching together" the $X(A)$s starting from any asset level forms an equilibrium path. When $X(A) = A$, it does.



(A) Constructing the interval $[S_1, S_2]$ (B) Double scaling

FIGURE 7. THRESHOLD FOR STRONG SELF-CONTROL.

defined S, and shows that our initial presumption is false. Therefore $X(A) < A$ for every A close enough to B. That establishes the existence of an initial range of assets for which a poverty trap is present, and so proves (i). Next, we work on:

(ii) There is $A_2 \geq A_1$ such that every $A \geq A_2$ exhibits strong self-control.

By nonuniformity, there is certainly some value of A for which $X(A) > A$. If the same inequality continues to hold for all $A' > A$, then by Proposition 2, strong self-control is established, not just at A but at every asset level beyond it. So the case that we need to worry about is one in which $X(A^*) \leq A^*$ for some asset level still higher than A. See Figure 7. Let S be the supremum asset level on $[B, A^*]$ for which $X(A) > A$; then as in the exposition for part (i), $X(S) = S$ and S is indeed sustainable.

By Lemma 1, the function $X(A)$ in the region $[B, S]$ can be scaled up and replicated as an equilibrium choice over the region $[S, S_1]$, where S_1 bears the same ratio to S as S does to B. Figure 7 depicts this in the dotted line with domain $[S, S_1]$. Note that there is a poverty trap near B, so the line lies below the 45^0 line to the right of B and to the right of S.

That said, there is one feature near S that cannot be replicated near B. Just to the right of S, it is possible to implement even smaller continuation assets by dipping into the zone to the left of S, and then accumulating upwards along $X(A)$

back towards S. Because these choices — shown by the solid line to the right of S in Figure 7 — favor current consumption over the future, they generate even lower equilibrium *values*, but they earn high enough *payoffs* so that they can be successfully implemented as equilibria. In turn, these lower values do a better job of forestalling impetuous deviations at still higher asset levels, and so it turns out that for asset levels close to S_1, the incentive constraints are relaxed and larger values of continuation assets (the solid line again) are sustainable. In particular, while S_1 is a sustainable asset level, it *also* permits accumulation: $X(S_1) > S_1$.

This argument creates a zone (possibly a small interval, but an interval nonetheless) just above S_1, call it $[S_1, S_2]$, over which (a) $X(A) > A$, and (b) each $S' \in [S_1, S_2]$ is sustainable. Part (a) follows from the fact that $X(S_1) > S_1$ and that X is nondecreasing. Part (b) follows from the fact that assets just to the right of S_1 were at least "almost sustainable" by virtue of the scaling argument of Lemma 1, and now must be deemed fully sustainable by virtue of the additional punishment properties that we've established in the region of S.

Panel B of Figure 7 now focusses fully on this zone and its implications. The following variation on Lemma 1 forms our central argument:

LEMMA **2.** *Suppose that asset levels S_1 and S_2 are both sustainable, and that $X(A) > A$ for all $A \in [S_1, S_2]$. Then there exists $\hat{A}$ such that $X(A) > A$ for all $A > \hat{A}$.*

The proof of the lemma is illustrated in the second panel of Figure 7. Define $\mu_i = \frac{S_i}{B}$ for $i = 1, 2$. Then for all positive integers k larger than some threshold K, the intervals $[\mu_1^k S_1, \mu_2^k S_2]$ and $[\mu_1^{k+1} S_1, \mu_2^{k+1} S_2]$ *must overlap*. It is easy to see why: $\mu_2^k S_2$ is just $\mu_2^{k+1} B$ while $\mu_1^{k+1} S_1$ is $\mu_1^{k+2} B$, and for large k it must be that μ_2^{k+1} exceeds μ_1^{k+1}.

Once this is settled, we can generate *any* asset level $A > \mu_1^K S_1$ by simply choosing an integer $k \geq K$, an integer m between 0 and k, and two values A_1 and A_2 from the interval $[S_1, S_2]$ so that

$$A = \mu_1^m A_1 \mu_2^{k-m}.$$

But each of A_1 and A_2 is sustainable, and moreoever, $X(A_i) > A_i$ for each of them, so that repeated application of Lemma 1 proves that $X(A) > A$. That proves Lemma 2.

But now the proof of the theorem is complete: by Proposition 2, if $X(A) > A$ for all A sufficiently large, the required threshold A_2 must exist.

6. Some Implications of the Theory

The main connection we emphasize in this paper is that there is a systematic link between credit limits and the ability to exercise self-control. The very same individual (psychologically speaking) can behave quite differently when he is close to the minimum level to which his assets can go. He has little power to exercise control over himself using internal rules, because the narrow margins below which he is constrained not to fall do not leave any room for suitable "punishments". In contrast, when that individual has sufficient wealth, he can use internal rules to carry out sustained accumulation, provided, of course, that his proclivity for current consumption is not extremely high.

It is evident that while our model is not scale-neutral, there is neutrality in a modified sense. One such sense (another is given in Lemma 1) is that the *ratio* of initial asset A to the lower bound B — fully determines an individual's ability to exercise self-control. We can rephrase all our observations in terms of this ratio. In particular, Proposition 4 can be interpreted as saying that there are two ratios μ_1 and μ_2, with $1 < \mu_1 \leq \mu_2 < \infty$, such that a poverty trap exists whenever $A/B < \mu_1$, while unlimited accumulation is possible whenever $A/B > \mu_2$.

6.1. **Ambiguous Effects of Changing Credit Limits.** A first implication of Proposition 4 is that an improvement in the credit limit has ambiguous effects, depending on initial assets A. Such an improvement lowers B. If that tips A/B over the threshold μ_2, sustained accumulation becomes possible where none was possible before. On the other hand, if A/B remains below μ_1 after the improvement, the individual will slide into an even deeper poverty trap.

6.2. **Asset-Specific Marginal Propensities to Consume.** A second implication is that the model naturally generates different marginal propensities to consume from income flows and assets. This phenomenon is studied in Hatsopoulos, Krugman and Poterba (1989), Thaler (1990) and Laibson (1997), though admittedly the empirical evidence for it may be somewhat debatable.[21] To see this, recall Section 2.1 and our interpretation there of the lower bound as some function of permanent income, presumably one that is related to the fraction of future labor income that can be seized in the event of a default. That is, if F_t stands for *financial* assets at date t and y for income at every date, then A_t is the present value of financial and labor assets:

$$A_t = F_t + \frac{\alpha y}{\alpha - 1},$$

while

$$B = \frac{\sigma \alpha y}{\alpha - 1}$$

for some $\lambda \in (0, 1)$. With this in mind, consider an increase in current financial assets F. Then B is unchanged, so that A/B must rise. Our proposition suggests that this will enhance self-control, so that accumulation is possible in a situation where previously it wasn't. In that case, the marginal propensity to consume out of an unforeseen change in financial assets could be "low".

In contrast, consider an equivalent jump in y, so that A rises by the same amount. Under our specification, B/y is constant so that A/B must *fall.* By the ratio interpretation of Proposition 4, self-control is damaged: the marginal propensity to consume from an unforeseen change in permanent income will be high. Indeed, even if B is a more complex increasing function of permanent income, it is only in the extreme case that B is entirely unchanged when y increases, and it is in that case that the propensities to consume out of the two asset classes will be the same. Otherwise, they will be different.

6.3. **External Versus Internal Commitments.** Our model is one that fully emphasizes internal rules for achieving self-control. An important extension is one to the case in which both internal and external commitments are available. The latter would include bank deposit schemes in which there are constraints on withdrawal, or legal commitments to make ongoing deposits (or both). Certainly, external commitments help when internal commitment fails, and this suggests that the asset-poor would have a higher demand for such arrangements.

[21]Some of the effects may be driven by variations in the different stochastic processes governing various sources of income.

Notice that external commitments effectively raise the value of the credit limit B, because they represent assets that cannot be drawn down. That suggests that unless *all* savings are carried out in the form of locked commitment schemes, such arrangements might damage other forms of "internal savings" as external assets accumulate. ("I have a retirement account, I don't need to save.")

That suggests that a judicious policy that takes advantage of external commitments below some asset threshold, coupled with a reliance on internal commitments when above that threshold, might be optimal. For instance, one might permit the agent to choose a particular savings targets, upon the attainment of which the lock-in is removed. To make this argument precise, we need some uncertainty in preferences or in the external environment which renders *exclusive* external arrangements infeasible.

6.4. **Who Wants External Commitments?** Finally, we comment on the economic characteristics of those individuals who might value external commitments. Clearly, these are the individuals who are asset-poor relative to their credit limit. The asset-rich would rather save on their own. The same observations would also be true of the income-poor and the income-rich provided B is unchanged across the two categories. On the other hand, these observations are reversed if B is a constant fraction of permanent income. In that case, and *controlling* for financial assets, it is the income-rich who would exhibit a greater desire for external commitment.

To be sure, the income-rich may also be asset-rich, so that the net effect is ambiguous. Nevertheless, the theory informs an empirical specification which can, in principle, be tested.

7. Numerical Analysis

8. Conclusions and Directions for Future Research

Proofs

Proof of Observation 1. Let $L(A)$ denote the infimum of all equilibrium values at A. By (3) and the restriction that $A_t \geq B$ for any feasible asset choice at date t, we have for any feasible consumption c_t at date t,

$$u(c_t) \geq u(\nu B),$$

so that $L(A) \geq (1-\delta)^{-1} u(\nu B) > -\infty$.

Suppose that $V \in \mathcal{V}(A)$. Let x be the prescribed initial asset choice, P the associated payoff. Then $P = (1-\beta)u\left(A - \frac{x}{\alpha}\right) + \beta V$, so by the equilibrium condition,

$$(1-\beta)u\left(A - \frac{x}{\alpha}\right) + \beta V \geq u\left(A - \frac{B}{\alpha}\right) + \beta\delta L(B),$$

and noting that $u\left(A - \frac{x}{\alpha}\right) \leq u\left(A - \frac{B}{\alpha}\right)$, we may conclude that

$$V \geq u\left(A - \frac{B}{\alpha}\right) + \delta L(B). \tag{12}$$

By applying (12) to $A = B$ and the value $L(B)$, or (if needed) a sequence of equilibrium values in $\mathcal{V}(B)$ that converge down to $L(B)$,

$$L(B) \geq u\left(B - \frac{B}{\alpha}\right) + \delta L(B). \tag{13}$$

Combining (12) and (13), the proof is complete. ■

Proof of Proposition 1. We claim that if $\mathcal{W}$ is nonempty, has closed graph, and satisfies (6), then it generates a correspondence with the same properties.

Let $\mathcal{W}'$ be the correspondence generated by $\mathcal{W}$. We first prove that $\mathcal{W}'$ is nonempty-valued.

Consider the function H on $[B, \infty)$ defined by

$$H(A) \equiv \max \mathcal{W}(A)$$

for all $A \geq B$. It is easy to see that H is usc. It follows that the problem

$$\max_{x \in [0, A/\alpha]} u\left(A - \frac{x}{\alpha}\right) + \beta\delta H(x)$$

is well-defined and admits a solution $x(A)$ for every $A \geq B$. Define

$$W \equiv u\left(A - \frac{x(A)}{\alpha}\right) + \delta H(x(A)).$$

Now observe that W satisfies (7) — pick $x = X(A)$ and $V = H(x(A))$. It also satisfies (8) — for each feasible x', take V' to be any element of $\mathcal{W}(x')$.

We next prove that $\mathcal{W}'$ has closed graph. Take any sequence (A_n, W_n) such that (i) $W_n \in \mathcal{W}'(A_n)$ for all n, and (ii) $(A_n, W_n) \to (A, W)$ as $n \to \infty$. We need to prove that $W \in \mathcal{W}'(A)$. For each n, pick continuation asset x_n and value $V_n \in \mathcal{W}(x_n)$ such that (7) is satisfied. Then for any limit point (x, V) of (x_n, V_n), we have $V \in \mathcal{W}(x)$, x feasible for A, and (7) also satisfied.

To verify (8) for (A, W), pick any x' feasible for A. Choose any sequence x'_n feasible for A_n such that $x'_n \to x'$. Because (8) holds for (A_n, W_n), there is $V'_n \in \mathcal{W}(x'_n)$ such that (8) holds. Let V' be any limit point of $\{V'_n\}$, then $V' \in \mathcal{W}(x')$, and so (8) holds for (A, W) at x'.

With this claim in hand, consider the iterated sequence $\{\mathcal{V}_k\}$. Because $\mathcal{V}_0$ is nonempty, has closed graph, and satisfies (6), so do all the $\mathcal{V}_k$'s. Moreover, for each $t \geq 0$, it is obvious that

$$\mathcal{V}_k(A) \supseteq \mathcal{V}_{k+1}(A)$$

for all $A \geq B$. Take infinite intersections and use nested compact sets (at each A) to argue that

$$\mathcal{V}(A) \equiv \bigcap_{t=0}^{\infty} \mathcal{V}_k(A)$$

is nonempty and compact-valued for every A.

Indeed, it is easy to see that $\mathcal{V}$ has compact graph on any compact interval $[B, D]$ of the domain,[22] and therefore it has closed graph everywhere.

By picking $V \in \mathcal{V}(A)$ and taking limits of continuation assets, values and punishments as $k \to \infty$, it is immediate that $\mathcal{V}$ generates itself and contains all other correspondences that do, so it is our equilibrium correspondence. ∎

In the light of Proposition 1, we may define $H(A)$ and $L(A)$ to be the maximum and minimum values of the equilibrium value correspondence. Because the graph of that correspondence is closed, it is easy to see that H is usc and L is lsc. Now

[22]On any compact interval, the (restricted) graphs of the $\mathcal{V}_k$'s are compact and their infinite intersection is precisely the graph of $\mathcal{V}$ restricted to the same interval. Therefore $\mathcal{V}$ has compact graph on this interval.

we can define the "deviation payoff" at any asset level A by

$$D(A) = \sup_{y \in [B, \alpha(1-\upsilon)A]} u\left(A - \frac{y}{\alpha}\right) + \beta\delta L(y).$$

It is easy to see that $D(A)$ is a nondecreasing function.

By a *path* from $A \geq B$ we will mean any feasible asset sequence $\{A_t\}$ with $A_0 = A$ and $B \leq A_{t+1} \leq \alpha(1-\upsilon)A_t$ for all $t \geq 0$. Each path generates a corresponding sequence of values $\{V_t\}$ in the obvious way: $V_t \equiv \sum_{s=t}^{\infty} \delta^{s-t} u\left(A_t - \frac{A_{t+1}}{\alpha}\right)$ for every $t \geq 0$.

LEMMA **3.** *A path $\{A_t\}$ is an equilibrium if and only if*

$$u\left(A_t - \frac{A_{t+1}}{\alpha}\right) + \beta\delta V_{t+1} \geq D(A_t) \tag{14}$$

for all t, where $\{V_t\}$ is the sequence of values generated by the path.

Proof. Sufficiency is a familiar consequence of the one-shot deviation principle. To prove necessity, note that if $\{A_t\}$ is an equilibrium path, then for every t,

$$u\left(A - \frac{A_{t+1}}{\alpha}\right) + \beta\delta V_{t+1} \geq \sup_{y \in [B, \alpha(1-\upsilon)A] \setminus A_{t+1}} u\left(A - \frac{y}{\alpha}\right) + \beta\delta L(y)$$

But $V_{t+1} \geq L(x)$ for each t, so this inequality can be extended to the condition

$$u\left(A - \frac{A_{t+1}}{\alpha}\right) + \beta\delta V_{t+1} \geq \sup_{y \in [B, \alpha(1-\upsilon)A]} u\left(A - \frac{y}{\alpha}\right) + \beta\delta L(y) = D(A).$$

■

Say that an asset level S is *sustainable* if there is a stationary equilibrium path from S, or equivalently (by Lemma 3) if

$$\left[1 + \frac{\beta\delta}{1-\delta}\right] u\left(\frac{\alpha - 1}{\alpha} A\right) \geq D(A).$$

The following lemma is central to the main line of argument in the paper.

LEMMA **4.** *Let $S > B$ be sustainable. Define $\mu = S/B > 1$. Then if $\{A_t\}$ is an equilibrium asset path from A_0, so is $\{\mu A_t\}$ starting from μA_0.*

Proof. Let ϕ be the equilibrium policy that sustains $\{A_t\}$ from A_0. Define a new policy function ψ as follows:

(i) For any h_t with $A_s \geq S$ for $s = 0, ..., t$, let $\psi(h_t) = \mu\phi\left(\frac{h_t}{\mu}\right)$, where the notation h_t/μ has the obvious meaning.

(ii) For any h_t with $A_s < S$ for some s, let k be the lowest such s. Define $h'_{t-k} = (A_k, ..., A_t)$. Let $\psi(h_t) = \phi_L(h'_{t-k})$, where ϕ_L is the equilibrium policy function that delivers the continuation value $L(A_k)$ starting from asset level A_k.

Consider any history h_t with $A_s \geq S$ for $s = 1, ..., t$. It is easy to check that the asset sequence generated through repeated application of ψ starting from h_t is the same as the asset sequence generated through repeated application of ϕ starting from $\frac{h_t}{\mu}$, with the latter scaled up by the factor μ. It follows that

$$P_\psi(h_t) = \mu^{1-\sigma} P_\phi\left(\frac{h_t}{\mu}\right) \tag{15}$$

and

$$V_\psi(h_t) = \mu^{1-\sigma} V_\phi\left(\frac{h_t}{\mu}\right). \tag{16}$$

To complete the proof, we show that ψ is an equilibrium.

Consider any h_t such that $A_s < S$ for some $s \leq t$, and let k be the first such period. Then by construction, as of period k the policy function ψ shifts to the equilibrium that delivers the value $L(A_k)$. So $\psi(h_t)$ is optimal given the continuation policy function.

Next consider any h_t such that $A_s \geq S$ for all $s \leq t$. Consider, first, any deviation to $A' \geq S$. Note that h_t/μ is a feasible history under the equilibrium ϕ, while the deviation to $(A'/\mu) \geq (S/\mu) = B$ is also feasible. It follows that

$$P_\phi\left(\frac{h_t}{\mu}\right) \geq u\left(\frac{A_t}{\mu} - \frac{A'}{\mu\alpha}\right) + \beta\delta V_\phi\left(\frac{h_t}{\mu}\right).$$

Multiplying through by $\mu^{1-\sigma}$ and using (15) and (16), we see that

$$P_\psi(h_t) \geq u\left(A_t - \frac{A'}{\alpha}\right) + \beta\delta V_\psi(h_t, A'), \tag{17}$$

which shows that no deviation to $A' \geq S$ can be profitable.

Now consider a deviation to $A' < S$. Because S is sustainable,

$$\left[1+\frac{\beta\delta}{1-\delta}\right]u\left(S\left(1-\frac{1}{\alpha}\right)\right)\geq u\left(S-\frac{A'}{\alpha}\right)+\beta\delta L(A'). \tag{18}$$

At the same time, a deviation from the going equilibrium to S is not profitable (by (17)), so

$$P_\psi(h_t)\geq u\left(A_t-\frac{S}{\alpha}\right)+\beta\delta V_\psi(h_t,S) \tag{19}$$

Using (16) along with $L(B)\geq \frac{1}{1-\delta}u\left(B\left(1-\frac{1}{\alpha}\right)\right)$ (see Observation 1), (19) becomes

$$\begin{aligned}
P_\psi(h_t) &\geq u\left(A_t-\frac{S}{\alpha}\right)+\beta\delta\mu^{1-\sigma}V_\phi\left(\frac{h_t}{\mu},B\right)\\
&\geq u\left(A_t-\frac{S}{\alpha}\right)+\beta\delta\mu^{1-\sigma}L(B)\\
&\geq u\left(A_t-\frac{S}{\alpha}\right)+\beta\delta\mu^{1-\sigma}\frac{1}{1-\delta}u\left(B\left(1-\frac{1}{\alpha}\right)\right)\\
&= u\left(A_t-\frac{S}{\alpha}\right)+\frac{\beta\delta}{1-\delta}u\left(S\left(1-\frac{1}{\alpha}\right)\right)\\
&= \left[u\left(A_t-\frac{S}{\alpha}\right)-u\left(S\left(1-\frac{1}{\alpha}\right)\right)\right]+\left[1+\frac{\beta\delta}{1-\delta}\right]u\left(S\left(1-\frac{1}{\alpha}\right)\right) \qquad (20)
\end{aligned}$$

Combining (18) and (20) and using the concavity of u (along with $S > A'$),

$$\begin{aligned}
P_\psi(h_t) &\geq \left[u\left(A_t-\frac{S}{\alpha}\right)-u\left(S\left(1-\frac{1}{\alpha}\right)\right)\right]+u\left(S-\frac{A'}{\alpha}\right)+\beta\delta L(A')\\
&= \left[u\left(A_t-\frac{S}{\alpha}\right)-u\left(S\left(1-\frac{1}{\alpha}\right)\right)\right]-\left[u\left(A_t-\frac{A'}{\alpha}\right)-u\left(S-\frac{A'}{\alpha}\right)\right]\\
&\quad +\left[u\left(A_t-\frac{A'}{\alpha}\right)+\beta\delta L(A')\right]\\
&\geq u\left(A_t-\frac{A'}{\alpha}\right)+\beta\delta L(A')\\
&= u\left(A_t-\frac{A'}{\alpha}\right)+\beta\delta V_\psi(h_t,A').
\end{aligned}$$

It follows that the deviation A' is unprofitable, so that ψ is an equilibrium. ■

Lemma **5.** *For any asset level A and any path $\{A_t\}$ from A with $A_t \leq A$ for all t,*

$$\frac{1}{1-\delta}u\left(\frac{\alpha-1}{\alpha}A\right) \geq \sum_{t=0}^{\infty}\delta^t u\left(A_t - \frac{A_{t+1}}{\alpha}\right). \tag{21}$$

Proof. Let Δ be the difference between the left and right hand sides in (21); then

$$\begin{aligned} \Delta &= \sum_{t=0}^{\infty}\delta^t\left[u\left(\frac{\alpha-1}{\alpha}A\right) - u\left(A_t - \frac{A_{t+1}}{\alpha}\right)\right] \\ &\geq u'\left(\frac{\alpha-1}{\alpha}A\right)\sum_{t=0}^{\infty}\delta^t\left[A - \frac{A}{\alpha} - A_t + \frac{A_{t+1}}{\alpha}\right] \\ &= u'\left(\frac{\alpha-1}{\alpha}A\right)\sum_{t=0}^{\infty}\delta^t\left[(A - A_t) - \frac{A - A_{t+1}}{\alpha}\right] \\ &= u'\left(\frac{\alpha-1}{\alpha}A\right)\left(\delta - \frac{1}{\alpha}\right)\sum_{t=0}^{\infty}\delta^t\left[A - A_{t+1}\right] \\ &\geq 0, \end{aligned}$$

where the first inequality uses the concavity of u. ■

As in the text, define $X(A)$ to be the largest equilibrium asset choice at A.

Lemma **6.** $X(A)$ *is well-defined, non-decreasing, and usc.*

Proof. $X(A)$ is the largest value of $A' \in [B, \alpha(1-\upsilon)A]$ satisfying

$$u\left(A - \frac{A'}{\alpha}\right) + \beta\delta H(A') \geq D(A) \geq u\left(A - \frac{y}{\alpha}\right) + \beta\delta L(y) \tag{22}$$

for all $y \in [B, \alpha(1-\upsilon)A]$. $X(A)$ is well-defined because H is usc. To show that $X(A)$ is non-decreasing, pick $A_1 < A_2$. If $u\left(A_2 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) \geq D(A_2)$, then we are done. Otherwise there is $x' \in [B, \alpha(1-\upsilon)A_2]$ such that

$$u\left(A_2 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) < u\left(A_2 - \frac{x'}{\alpha}\right) + \beta\delta L(x'), \tag{23}$$

which implies

$$u\left(A_2 - \frac{x'}{\alpha}\right) - u\left(A_2 - \frac{X(A_1)}{\alpha}\right) > \beta\delta\left[H(X(A_1)) - L(x')\right] \tag{24}$$

There are two cases to consider: (i) $x' \leq X(A_1)$, and (ii) $x' > X(A_1)$. In case (i), x' is feasible under A_1, so that

$$(25) \qquad u\left(A_1 - \frac{x'}{\alpha}\right) - u\left(A_1 - \frac{X(A_1)}{\alpha}\right) \leq \beta\delta\left[H(X(A_1)) - L(x')\right]$$

But (24) and (25) together contradict the concavity of u.

In case (ii), we can combine (22) and (23) to see that

$$(26) \qquad \begin{aligned} u\left(A_2 - \frac{x'}{\alpha}\right) + \beta\delta L(x') &> u(A_2 - X(A_1)) + \beta\delta H(X(A_1)) \\ &\geq u(A_2 - y) + \beta\delta L(y) \end{aligned}$$

for all $y \leq X(A_1)$. We now construct an equilibrium starting from A_2 as follows: any choice $A < X(A_1)$ is followed by the continuation equilibrium generating $L(A)$, and any choice $A \geq X(A_1)$ is followed by the continuation equilibrium generating $H(A)$. Because H is usc, there exists some z^* that maximizes $u\left(A_2 - \frac{z}{\alpha}\right) + \beta\delta H(z)$ on $[X(A_1), \alpha(1-v)A_2]$; in light of (26) and the fact that $u\left(A_2 - \frac{x}{\alpha}\right) + \beta\delta H(x) \geq u\left(A_2 - \frac{x}{\alpha}\right) + \beta\delta L(x)$, all choices $A < X(A_1)$ are strictly inferior to z^*. Thus z^* is an equilibrium choice at A_2, so that $X(A_2) \geq z^* \geq X(A_1)$.

Finally, we show that X is usc. Choose any asset level A^*. Certainly, $\lim_{A\uparrow A^*} X(A) \leq X(A^*)$ because $X(A)$ is nondecreasing. Now consider any decreasing sequence $A^k \downarrow A^*$, and let X^* denote the limit of $X(A^k)$ (this is well-defined because $X(A)$ is non-decreasing). For each k, we have $u\left(A^k - \frac{X(A^k)}{\alpha}\right) + \beta\delta H(X(A^k)) \geq D(A^k)$. Taking limits and using the fact that H is usc, we have $u\left(A^* - \frac{X^*}{\alpha}\right) + \beta\delta H(X^*) \geq \lim_{k\to\infty} D(A^k)$. Because $D(A)$ is nondecreasing, it follows that $u\left(A^* - \frac{X^*}{\alpha}\right) + \beta\delta H(X^*) \geq D(A^*)$, which in turn implies $X(A^*) \geq X^* = \lim_{A\downarrow A^*} X(A)$. In fact, because $X(A)$ is non-decreasing, we must have $X(A^*) = \lim_{A\downarrow A^*} X(A)$. ■

LEMMA **7.** *If $X(A) = A$, then A is sustainable.*

Proof. Let $\{A_t\}$ be an equilibrium path from A that supports $X(A) = A$; i.e., with $A_1 = A$. By Lemma 3,

$$u\left(\frac{\alpha - 1}{\alpha}A\right) + \beta\delta V_1 \geq D(A).$$

By Lemma 5, $V_1 \leq (1-\delta)^{-1}u\left(\frac{\alpha-1}{\alpha}A\right)$. Using this in the inequality above, we must conclude that A is sustainable. ■

Lemma 8. *In the nonuniform case,*

$$\frac{\beta\delta}{1-\delta}(\alpha-1)<1. \tag{27}$$

Proof. We claim that if $\frac{\beta\delta}{1-\delta}(\alpha-1)\geq 1$, then there exists a linear Markov equilibrium policy function $\phi(A)=kA$ with $k>1$, which implies uniformity.

To this end, assume that the individual will employ the policy function $\phi(A)=kA$ with $k\in[1,\alpha]$ for all $A\geq B$. In that case, the the individual's current problem is to solve

$$\max_{x\in[B,\alpha(1-\upsilon)A]}\frac{1}{1-\sigma}\left[\left(A-\frac{x}{\alpha}\right)^{1-\sigma}+\beta\delta Q x^{1-\sigma}\right]$$

where

$$Q\equiv\frac{(\alpha-k)^{1-\sigma}}{\alpha^{1-\sigma}\left(1-\delta k^{1-\sigma}\right)} \tag{28}$$

The corresponding first-order condition is:

$$\frac{1}{\alpha}\left(A-\frac{x}{\alpha}\right)^{-\sigma}=\beta\delta Q x^{-\sigma}$$

After some manipulation, we obtain

$$\frac{A}{x}=\frac{1}{\alpha}+\left(\frac{1}{\alpha\beta\delta Q}\right)^{1/\sigma}\equiv\frac{1}{k^{*}} \tag{29}$$

Note that $x=k^{*}A$. Accordingly, the policy function is an equilibrium if $k^{*}=k$. Substituting (28) into (29) and rearranging yields

$$k^{\sigma}=\alpha\beta\delta+(1-\beta)\,\delta k \tag{30}$$

Define $\Lambda(k)\equiv k^{1-\rho}$ and $\Phi(k)=\alpha\beta\delta+(1-\beta)\,\delta k$. Notice that $\Lambda(1)\leq\Phi(1)$ (given that $\frac{\beta\delta}{1-\delta}(\alpha-1)\geq 1$), and $\Lambda(\alpha)>\Phi(\alpha)$ (given the transversality condition, $\delta\alpha^{1-\sigma}<1$). By continuity, it follows that there exists a solution on the interval $[1,\alpha)$, which establishes the claim and hence the lemma. ■

Proof of Proposition 4, part (i). We first claim that there is no interval of the form $[B,B+\epsilon]$ with $X(A)>A$ on that interval. For suppose there were. By nonuniformity, $X(A')<A'$ for some $A'>B+\epsilon$. Because X is nondecreasing and usc, there is $S>B$ and $\epsilon'>0$ such that $X(S)=S$ and such that $X(A')<A'$ for

all $A' \in (S, S+\epsilon')$. By Lemma 7, S is sustainable. Defining $\mu \equiv S/B$ and invoking Lemma 4, we see that the continuation asset $\mu X(A'/\mu)$ is an equilibrium choice for every $A' \in [S, S+\mu\epsilon]$. But then $X(A') \geq \mu X(A'/\mu) > A'$, a contradiction.

In particular, $X(B) = 0$, and for all $\epsilon > 0$, there exists $A^L_\epsilon \in (B, \epsilon)$ such that $X(A^L_\epsilon) < A^L_\epsilon$.

Now, if part (i) of the proposition is false, then for all $\zeta > 0$ there also exists $A^H_\zeta \in (B, \zeta)$ such that $X(A^H_\zeta) \geq A^H_\zeta$. Taking $\zeta = A^L_\epsilon - B$ and recalling that $X(A)$ is non-decreasing (Lemma 6), we see that there must also exist $S_\epsilon \in (B, \epsilon)$ (specifically, lying between A^H_ζ and A^L_ϵ) such that $X(S_\epsilon) = S_\epsilon$.

Now consider the following artificial maximization problem, for any $A > B$:

$$\max_{A' \in [0,A]} u\left(A - \frac{A'}{\alpha}\right) + \frac{\beta\delta}{1-\delta} u\left(A'\left(1 - \frac{1}{\alpha}\right)\right)$$

It is easy to check that the solution is

$$A' = \gamma A$$

where

$$\gamma = \frac{\alpha}{1 + \xi^{-\frac{1}{\sigma}}(\alpha - 1)}$$

and

$$\xi = \frac{\beta\delta}{1-\delta}(\alpha - 1)$$

Given nonuniformity and Lemma 8, we know that $\xi < 1$, which (given $\sigma > 0$) implies $\gamma < 1$. So applying the maximization problem to the asset level S_ϵ, we see that for ϵ sufficiently small, $\gamma S_\epsilon < B < S_\epsilon$. Because u is concave, this implies that

$$u\left(S_\epsilon - \frac{B}{\alpha}\right) + \frac{\beta\delta}{1-\delta} u\left(B\left(1 - \frac{1}{\alpha}\right)\right) > (1 + \frac{\beta\delta}{1-\delta}) u\left(S_\epsilon\left(1 - \frac{1}{\alpha}\right)\right). \tag{31}$$

Because $X(S_\epsilon) = S_\epsilon$, Lemma 7 tell us that S_ϵ is sustainable. The equilibrium payoff from sustaining S_ϵ is the RHS of (31). On the other hand, because $X(B) = B$, the payoff from a deviation that sets continuation assets equal to B is the LHS of (31). That means that (31) is inconsistent with the existence of an equilibrium that maintains S_ϵ forever, a contradiction. ∎