POVERTY AND SELF-CONTROL

B. Douglas Bernheim
Stanford University and NBER

Debraj Ray
New York University

Sevin Yeltekin
Carnegie Mellon University

January 2013

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

ABSTRACT

The absence of self-control is often viewed as an important correlate of persistent poverty. Using a standard intertemporal allocation problem with credit constraints faced by an individual with quasi-hyperbolic preferences, we argue that poverty damages the ability to exercise self-control. Our theory invokes George Ainslie's notion of "personal rules," interpreted as subgame-perfect equilibria of an intrapersonal game played by a time-inconsistent decision maker. Our main result pertains to situations in which the individual is neither so patient that accumulation is possible from every asset level, nor so impatient that decumulation is unavoidable from every asset level. Such cases always possess a threshold level of assets above which personal rules support unbounded accumulation, and a second threshold level of assets below which there is a "poverty trap": no personal rule permits the individual to avoid depleting all liquid wealth. In short, poverty perpetuates itself by undermining the ability to exercise self-control. Thus policies designed to help the poor accumulate assets may be highly effective, even if they are temporary. We also explore the implications for saving with easier access to credit, the demand for commitment devices, the design of accounts to promote saving, and the variation of the marginal propensity to consume across classes of resource claims.

Bernheim's research was supported by National Science Foundation Grants SES-0752854 and SES-1156263. Ray's research was supported by National Science Foundation Grant SES-0962124.

1. INTRODUCTION

The absence of self-control is often viewed as an important correlate 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 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, 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

---

[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* ("Leave Well Alone," November 18, 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 Udry (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 (2011) cite other studies that also show high rates of return to small investments.

[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, and Ashraf, Karlan, and Yin (2006) on the use of a commitment savings product in the Philippines. Aliber (2001), Gugerty (2007) and Anderson and Baland (2002) view ROSCA participation as a commitment device; see also the theoretical contributions of Ambec and Treich (2007) and Basu (2011). Duflo, Kremer, and Robinson (2010) explain fertilizer use (or the lack of it) in Kenya as a lack of commitment. In the ongoing debate on whether to overhaul the public distribution system for food in India to an entirely cash-based program, individual commitment issues figure prominently; see Khera (2011).

[4]See, for instance, Ainslie (1975, 1992), Thaler and Shefrin (1981), Akerlof (1991), Laibson (1997), O'Donoghue and Rabin (1999) or Ashraf, Karlan and Yin (2006). 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).

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 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 Ainslie (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"). It is natural to model such 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 Ainslie'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.

[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 *United Nations Development Program* Regional Report for Latin America, 2010, which implements these ideas. However, this complementary approach does not generate a demand for commitment devices.

[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 Ainslie's theory.

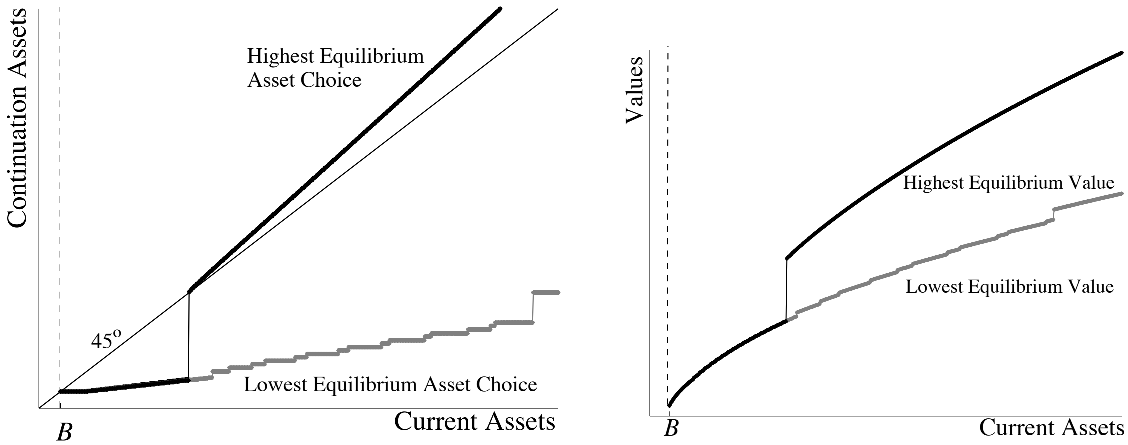
We assume that time-inconsistency arises from *quasi-hyperbolic discounting* (also known as $\beta\delta$-discounting), a standard model of intertemporal preferences popularized by Laibson (1994, 1996, 1997) and O'Donoghue and Rabin (1999). To determine the full scope for self-control, we study the set of *all* subgame-perfect Nash equilibria. To avoid excluding any viable personal rules, we impose no restrictions whatsoever on strategies (such as stationarity, or the use of Markov punishments). 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 self-control (through the use of sustainable personal rules) to accumulate greater wealth, and when she cannot.[9] 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 depleted 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 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 utility functions are homothetic, so we deliberately eliminate any preference-based relationship between assets and savings. (We return to this point when connecting our model to related literature.)

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

[8]Exceptions include Laibson (1994), Bernheim, Ray, and Yeltekin (1999), and Benhabib and Bisin (2001).
[9]Our distinctive focus on personal rules as history-dependent strategies can, of course, be questioned on the grounds that human life-spans are in fact finite, causing such rules to unravel from the terminal period. That criticism is not specific to our model, but applies to all analyses of infinite horizon games. That literature offers a number of potential answers; e.g., the unravelling logic can be overturned by examining epsilon-equilibria in finite horizon games (Fudenberg and Levine, 1983), introducing multiple stage-game equilibria in finite horizon games (Benoîit and Krishna, 1985), or by studying games in which the probability of continuation declines to (but does not reach) zero over time (Bernheim and Dasgupta, 1996).



(A) Equilibrium Asset Choices (B) Equilibrium Values

FIGURE 1. ACCUMULATION AND VALUES AT DIFFERENT ASSET LEVELS.

question by solving the model numerically using standard tools. (For a complete explanation of our computational methods, and for details on all computational examples presented in the text, see the Appendix.) Figure 1 illustrates the results of one such exercise.[10] The horizontal axes in each panel measure assets in the current period. The vertical axis in panel (A) similarly measures continuation asset choices for 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; see panel (A). Starting 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.[11]

[10]For this exercise, we set the rate of return equal to 30%, the discount factor equal to 0.8, the hyperbolic parameter (β) equal to 0.4, and the constant elasticity parameter of the utility function equal to 0.5. We chose these values so that the interesting features of the equilibrium set are easily visible; qualitatively similar features arise for more realistic parameter values.

[11]This is a more subtle point that cannot be seen directly from Figure 1, though it is indicative. The reason it is more subtle is that repeated application of the highest continuation asset need not be an equilibrium, and moreover, even if it were, it need not be the most attractive equilibrium.

The example motivates 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 limited 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. This is shown in panel (B) of the Figure, which plots equilibrium values. By a standard recursive argument, the lowest equilibrium value serves as the worst punishment, but notice that the lowest value jumps upwards; indeed, it does so at several asset levels.[12] 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 always an asset level below which liquid wealth is exhausted in finite time (that is, there exists a poverty

[12]The jagged nature of the lowest value in panel (B) is not a numerical artifact; it reflects actual jumps.

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:[13] 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.[14]

Our analysis has a number of provocative implications for economic behavior and public policy. We highlight five. 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. 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. But there is an important qualification: those who fail to make the transition fall further into debt.

Third, our analysis suggests a particular pattern of demand for precommitment devices (such as retirement accounts or fixed deposit schemes) as a function of wealth. In general, considerations of flexibility dictate that full precommitment is neither possible nor desirable. So people must rely to some extent on internal mechanisms for self-control, while seeking some form of supplementary external commitment mechanism. But the use of external commitments may undermine the efficacy of internal mechanisms by rendering personal rules ineffective. That isn't an issue when those personal rules are ineffective to begin with, so there should be a high demand for external commitment

[13]Though there is a 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.

[14]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 suitably normalized payoffs are often nonmonotonic. Also, identifying Markov equilibria is more computationally challenging than determining the key features of subgame-perfect equilibria.

devices in such cases (e.g., by low-wealth individuals). But other individuals will avoid the opportunity to lock up funds, even when they wish to save, because the lock-up moderates the consequences of a lapse in discipline, thereby making self-control more difficult to sustain. Presumably, these are individuals with assets already beyond the viable threshold.

Fourth, our analysis has implications for the design of programs intended to stimulate saving by providing access to special accounts (e.g., for retirement, education, home purchase, or other purposes). Virtually all such programs entail commitments, but the nature of those commitments differs considerably across programs. Based on our analysis, a particularly attractive design would require the individual to establish a target and lock up all funds until the target is achieved, at which point the lock is removed and all funds become liquid. Pilot programs with such features have indeed been tested in developing countries.[15] Notice how this argument follows by essentially applying the preceding observation to different levels of assets as they are endogenously accumulated.

Finally, our analysis provides a potential explanation for the observation that the marginal propensity to consume differs across classes of resource claims. In particular, the MPC from an unforeseen increase in permanent income may be relatively high because that development erodes self-control. Accordingly, our theory provides a new perspective on the excess sensitivity of consumption to income.

As noted above, we build on our unpublished working paper (Bernheim, Ray, and Yeltekin (1999)), which made its points through simulations, but did not contain analytical results. Our questions are related to those of Banerjee and Mullainathan (2010), who also argue that self-control problems give rise to low asset traps. Though their objective is similar, the 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 iPads, say) then becomes an empirical matter. But we avoid such hardwiring entirely by studying

[15]See Ashraf, Karlan, and Yin (2006), as well as Karlan, McConnell, Mullainathan, and Zinman (2010).

homothetic preferences in an established model of time-inconsistency. The phenomena we study are traceable to a single built-in feature: 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.[16]

The rest of the paper is organized as follows. Section 2 describes the model and definition of equilibrium. Section 3 introduces the set of equilibrium values and provides a characterization of that set. Section 4 defines self control, and Section 5 studies the relationship between self-control and the initial level of wealth. Section 6 describes additional implications of the theory. Section 7 presents conclusions and some directions for future research. Proofs are collected in Section 8. An Appendix describes our computational methods, as well as details for all numerical examples.

2. Model

2.1. **Feasible Set and Preferences.** The feasible set links current assets, current consumption and future assets, starting from an initial asset level A_0:

$$c_t = A_t - (A_{t+1}/\alpha) \geq 0, \tag{1}$$

and, in addition, imposes a lower bound on assets

$$A_t \geq B > 0. \tag{2}$$

Our leading interpretation of the lower bound B is that it is a credit constraint.[17] For instance, if F_t stands for *financial* wealth at date t and y for income at each date, then

[16]Our model is also related to Laibson (1994) and Benhabib and Bisin (2001), except for the all-important difference of an imperfect credit market. 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 below, that every equilibrium path can be suitably scaled to all levels of initial assets, so that there is no relationship between poverty and self-control.

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

A_t is the present value of financial 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 analyze; see Laibson (1994). On the other hand, if she can borrow only some fraction $(1 - \lambda)$ of lifetime income, then $B = \lambda\alpha y/(\alpha - 1)$.

Individuals have quasi-hyperbolic preferences: 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 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. Put another way, every scale effect in this paper 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 asset sequence $\{A_t\}$ 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. This program is well-defined provided utilities do not diverge, for which we assume that

$$\gamma \equiv \delta^{1/\sigma}\alpha^{(1-\sigma)/\sigma} < 1. \tag{3}$$

We presume throughout that the Ramsey program exhibits growth, which imposes

$$\delta\alpha > 1. \tag{4}$$

Under (3) and (4), the value $R(A)$ of the Ramsey program is finite, and

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

while assets grow exponentially:

$$A_{t+1} = A_0 \left(\delta^{1/\sigma}\alpha^{1/\sigma}\right)^t = A_0 \left(\gamma\alpha\right)^t.$$

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 vanishingly small consumption levels (see Laibson (1994) for a discussion of this point). We find such punishments rather contrived and unrealistic, and eliminate them by assuming that consumption is bounded below at every asset level. More precisely, we assume that at every date,

$$c_t \geq \upsilon A_t, \tag{5}$$

where υ is to be thought of as a small but positive number. It is formally enough to presume that $\upsilon < 1-\gamma$, so that Ramsey accumulation can occur unhindered, but the reader is free to think of this bound as tiny. Notice that we take the lower bound on consumption to be proportional to assets so as to avoid introducing an artificial scale effect through this constraint.

2.3. **Equilibrium.** A choice of continuation asset A' is *feasible* given A, if $A' \in [B, \alpha(1-\upsilon)A]$. A *path* is any sequence of assets with A_{t+1} feasible given A_t; so (1), (2) and (5) are satisfied. A *history* h_t at date t is a "truncated path" of assets $(A_0, \ldots, A_t)$ up to date t. Write $A(h_t) = A_t$ for the asset level at the start of date t following history h_t. A *policy* ϕ specifies a continuation asset $\phi(h_t)$ following every history, which must be feasible given $A(h_t)$. If h_t is a history and x a feasible asset choice, denote by $h_t.x$ the subsequent history generated by this choice. 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_ϕ by

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

for every history h_t. Values exclude the hyperbolic factor β, while payoffs include them.

An *equilibrium* is a policy such that at every history h_t and x feasible given $A(h_t)$,

$$P_\phi(h_t) \geq u\left(A(h_t) - \frac{x}{\alpha}\right) + \delta\beta V_\phi(h_t.x). \tag{6}$$

That is, an equilibrium may be viewed as the assignment of a continuation value for every choice of continuation asset (at any given history), where the *actual* continuation asset at that history is taken to be the one that maximizes the right hand side of (6) over all these specifications. For some of our observations, it will be useful to presume that a convex set of equilibrium continuation values is available at every asset level. We therefore suppose that following any asset choice, continuation values can be chosen (if needed) using a public randomization device.[18] Equivalently, an asset choice in period t is followed by a lottery over continuation plans starting in period t+1. That implies an obvious enlargement of the notion of a policy, the details of which we skip here.

3. Existence and Characterization of Equilibrium

For each $A \geq B$, let $\mathcal{V}(A)$ be the set of all equilibrium values available at 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 (3)) that

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

where $R(A)$ is the Ramsey value. Once (5) rules out unrealistic Ponzi-like cascades that generate arbitrarily low utility, a tighter bound is available for worst values:

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

$$L(A) \geq u\left(A - \frac{B}{\alpha}\right) + \delta L(B) \geq u\left(A - \frac{B}{\alpha}\right) + \frac{\delta}{1-\delta}u\left(\frac{\alpha-1}{\alpha}B\right) \tag{7}$$

Notice how Observation 1 kicks in as long as we place *any* (small) lower bound on consumption, as described in (5). 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 $A \geq B$,

$$\mathcal{W}(A) \subseteq \left[u\left(A - \frac{B}{\alpha}\right) + \frac{\delta}{1-\delta}u\left(\frac{\alpha-1}{\alpha}B\right), R(A)\right]. \tag{8}$$

[18] Here, "public" randomization simply means that, in each period, the individual observes the realization of a random variable, and does not subsequently forget it.

Say that $\mathcal{W}$ *supports the value* w *at asset level* A if there is a feasible asset choice x and $V \in \mathcal{W}(x)$ — a *continuation* $\{x, V\}$ in short — with

$$w = u\left(A - \frac{x}{\alpha}\right) + \delta V, \tag{9}$$

while for every feasible x',

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

for some $V' \in \mathcal{W}(x')$. That is, the value w at A can be created in an "incentive-compatible way" by choosing continuation values from $\mathcal{W}$. Now say that $\mathcal{W}$ *generates* the correspondence $\mathcal{W}'$ if for every $A \geq B$, $\mathcal{W}'(A)$ is the convex hull of all values supported at A by $\mathcal{W}$. Notice how the use of the convex hull captures public randomization (in the sense that an asset choice can yield a lottery over continuation values).

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.

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* $A \geq B$,

$$\mathcal{V}(A) = \bigcap_{k=0}^{\infty} \mathcal{V}_k(A). \tag{11}$$

Also, $\mathcal{V}$ *is convex-valued and has closed graph.*

This proposition is useful in that it establishes existence of equilibrium, though the method used may not apply more generally to all games with state variables.[19] The "generation logic" of the proposition inspires algorithms for numerical calculations along well-known lines, which we employ in all the exercises; see Appendix for details.[20]

[19] For more general existence theorems, see Goldman (1980) and Harris (1985).

[20] Incidentally note that the closed-graph property does not follow from a standard nested compact sets argument, because the sets in question (the graphs of $\mathcal{V}_k$) are not compact. It should also be noted that

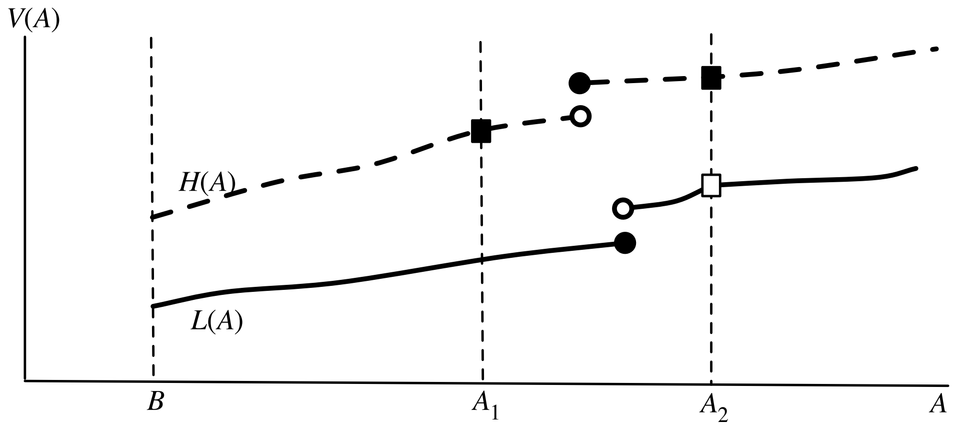


FIGURE 2. EQUILIBRIUM VALUES.

Figure 2 illustrates equilibrium values. Imagine supporting the highest value $H(A_1)$ at asset level A_1. That might require the choice of A_2 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 (6) 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 $x \neq A_2$.

4. SELF CONTROL

Viewed in the spirit of Ainslie's definition, the possibility of self-control via a sustainable personal rule refers to a feature of *some* element of the equilibrium correspondence. One might, for instance, say that self-control is possible if the Ramsey outcome itself is an equilibrium. That definition would require, of course, that the agent entirely transcend her hyperbolic urges. All other attempts, including accumulation at rates close to the Ramsey path, must then be deemed a failure of self-control, which we find too strong. We therefore employ a weak definition: 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 also be interested in whether the individual is capable of indefinite accumulation. Say that there is *strong self-control* at A if the agent is capable of unbounded accumulation — i.e., $A_t \to \infty$ — along some equilibrium path from A.

public randomization is not needed to establish existence; the same argument would work without it, except that $\mathcal{V}$ would not generally be convex-valued.

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. We 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* 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 study perfect credit markets in this paper, the observation is worth noting: if continuation asset x can be sustained at asset level A, then continuation asset λx can be sustained when the asset level is λA, for any $\lambda > 0$. In particular, if self-control is possible at some asset level, it is possible at all levels. Indeed, we've deliberately constructed the model in this fashion, so to understand better the scale effects created by introducing imperfect credit markets.

Therefore, the *nonuniform* cases are of primary interest to us. In these cases, 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 just when a case is nonuniform, such cases exist in abundance (we confirm this by numerical analysis).

We close this section with an intuitive yet nontrivial characterization of self control. Consider the *largest continuation asset*: the highest 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 self-control 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) are obvious, but parts (ii) and (iii), while intuitive, need a more extensive argument. Part (iii) will follow from the additional observation that X is nondecreasing and usc. Part (ii) will need more work to prove. Yet, if we take the proposition on faith for now, it will help us in visualizing the proof of the main theorem. It is worth mentioning that, under the conditions of part (ii), the value-maximizing equilibrium involves unbounded accumulation. That is noteworthy because value-maximization may be regarded as the most attractive from a long-run welfare perspective.[21]

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 rather simplistic failure of neutrality opens the door to all sorts of more interesting failures. 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.

These internal checks and balances are not merely technical, but descriptive (we feel) of individual ways of coping with commitment problems. One coping mechanism is

[21]See the "long-run" criterion for quasi-hyperbolic discounting in Bernheim and Rangel (2009).

"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 engaging in a behavior shift; for instance, she might go on a temporary consumption spree. In our theory, such a binge must be a valid continuation equilibrium. The threat of a "credible binge" might then help to keep the agent in check.

With this "internal mechanism" in mind, let's ask why an abundance of assets might help an individual to exhibit self-control. The ability to exercise control must depend on the severity of the consequences following an impetuous act of 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.[22]

Indeed, in the context of our model, the "severity of punishment" isn't monotonic in assets. Recall Figure 1 in the Introduction, which makes this point. Panel (B) displays highest and lowest value selections from the equilibrium correspondence. The lowest selection is $L(A)$. It jumps several times, showing that in general, punishment values (even after deflating by higher asset values) cannot be monotonically 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. We return to this point after we explain the simple structure of worst punishments in this model.

5.1. **Worst Punishments.** We will show that worst punishments involve a single spell of "excessive" expenditure, followed by a return to (approximately) the best possible continuation value. To formally describe this property, define, for any $A > B$, $H^{-}(A)$ by the left limit of $H(A^n)$ as $A^n \uparrow A$, with $A^n < A$ for all n. This is a well-defined concept because H is nondecreasing and therefore possesses limits from the left.

[22]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)).

PROPOSITION **3.** *The worst equilibrium value at any asset level A is implemented by choosing the smallest possible continuation asset at A; call it Y. Moreover, if $Y > B$, the associated continuation value V satisfies $V \geq H^-(Y)$.*

The proof is simple and instructive enough to be included in the main text.

Proof. Let Y be the smallest equilibrium choice of continuation asset at A, with associated continuation value V. Then the following natural no-deviation constraint applies:

$$u\left(A - \frac{Y}{\alpha}\right) + \beta\delta V \geq D(A), \tag{12}$$

where $D(A)$ is the supremum of all "deviation payoffs," in each of which every deviation to an alternative asset choice is "punished" by the lowest equilibrium value available at that asset.[23] If (12) is slack, it is easy to show that Y must equal B and that V can be set equal to $L(B)$.[24] That generates the lowest possible equilibrium value at A and there is nothing left to prove; see the first inequality in Observation 1.

Otherwise (12) is binding for Y. In this case,

$$u\left(A - \frac{Y}{\alpha}\right) + \beta\delta V = D(A) \leq u\left(A - \frac{A'}{\alpha}\right) + \beta\delta V'. \tag{13}$$

for any other equilibrium continuation $\{A', V'\}$ at A. Because $A' \geq Y$ by definition, (13) shows that $V' \geq V$. It follows that

$$u\left(A - \frac{Y}{\alpha}\right) + \delta V \leq u\left(A - \frac{A'}{\alpha}\right) + \delta V', \tag{14}$$

so that once again, $\{Y, V\}$ implements minimum value at A.

To complete the proof, suppose that $Y > B$ while at the same time, $V < H^-(Y)$. Then it is obviously possible to reduce Y slightly while increasing continuation value at the same time.[25] Moreover, the new continuation has higher payoff, so it must be supportable as an equilibrium. Yet it has a lower continuation asset, which contradicts the definition of Y. ∎

[23] The function $D(A)$ is formally defined in Section 8, where we deal with various technicalities arising from lack of the continuity in the value correspondence; see equation (20). Note that Lemma 3 following that equation establishes (12).
[24] For details, see Footnote 51 in Section 8.
[25] Because $V < H^-(Y)$, there exists $Y' < Y$ and $V' \in \mathcal{V}(Y)$ such that $V' > H^-(Y)$.

The heart of the argument is (14). If two continuations have the same payoff, the one that exhibits the larger upfront consumption must have the lower *value*. Payoffs include the factor β, which devalues future consumption. When β is "removed," as it is in the computation of value, the continuation with higher consumption today has lower value. That is why worst punishments exhibit a large binge to begin with; in fact, the largest possible credible binge. The binge is then followed by a reversion to the best possible equilibrium value — or approximately so, in a sense made precise in the proposition.[26]

Two more remarks are worth noting about lowest values, or optimal punishments. First, the associated actions have an extremely simple and plausible structure. No unrealistically complex rules are followed that might justify a restriction to "simpler" notions, such as Markov punishments. 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. Second, there is a sense in which these punishments are reasonably immune to renegotiation. While the earlier, deviating self fears the low-value 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.

Finally, while optimal punishments are reminiscent of the carrot-and-stick property for optimal penal codes in repeated games (Abreu (1988)), there is no reason why that property should hold, in general, for games with state variables, of which our model is an example. In this model, the particular structure arises 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. 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 repeated games.

5.2. **The Relationship Between Wealth and Self-Control.** The argument used to establish Proposition 3 is also informative on the issue of "jumps" in worst punishments.

[26]We note again that reversion to the best continuation value occurs, 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^{-}.

Suppose that the continuation $\{Y, V\}$ supports the lowest value at A. Let d be a "best deviation" choice of asset at A; namely, a choice of asset that attains the highest deviation payoff $D(A)$. To make our point, suppose that the no-deviation constraint is binding for the continuation $\{Y, V\}$, as it typically will be. Then

$$u\left(A - \frac{Y}{\alpha}\right) + \beta\delta V = u\left(A - \frac{d}{\alpha}\right) + \beta\delta L(d), \tag{15}$$

and the path associated with the initial choice of d is therefore also an equilibrium path. Recall from Proposition 3 that Y is the lowest possible choice of continuation asset at A. That proves that $d \geq Y$. For if d were smaller, then by the same argument employed in Section 5.1, the equilibrium *value* associated with d would be even smaller than that associated with $\{Y, V\}$, a contradiction.

So d is no smaller than Y, and in general will strictly exceed Y. It is precisely then that jumps can occur. To see this, increase A. Because $d > Y$, the strict concavity of utility forces the right hand side of the no-deviation constraint (15) to go up faster than the left hand side. To maintain that constraint, Y and V will need to change. But — depending on the shape of the equilibrium value correspondence — no *local* adjustment might suffice: the change may well have to be discrete. That will lead to an upward jump in L.[27] Numerical analysis tells us that such a scenario is chronic.

The possibility that worst equilibrium values can abruptly rise with wealth leads to the nihilistic suspicion that there is no general connection between wealth and self-control. Nevertheless, not one of the numerical examples that we have studied bears out this suspicion. Bernheim, Ray and Yeltekin (1999) find through simulations that either we are in one of the two uniform cases (accumulation possible everywhere, or accumulation impossible anywhere), or the situation looks like Figure 1. 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:*

[27] More generally, the constraint set is not continuous in A, leading to a failure of the well-known "maximum theorem."

(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 through any sustainable personal rule, and they must deplete all their wealth. Yet 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 is 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.

Notice that 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, a "single crossing" of the highest asset choice $X(A)$ over the 45^0 line may not be guaranteed, at least under the assumptions that we have made so far.[28] From this perspective, the fact that after a finite threshold all such 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, imperfect credit markets destroy scale-neutrality in our theory. (The constant elasticity of preferences assures us that otherwise, the model would be fully scale-neutral.) Yet variations of scale-neutrality survive. One variation that is particularly germane to our argument is given in Observation 2 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.[29]

[28]We have neither been able to rule out multiple crossings nor to find a numerical example with multiple crossings.

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

Observation **2.** *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.*

To understand this result, first think of S as a new lower bound on assets. Then the constant elasticity of utility together with linearity in the rate of return to assets together guarantee that any equilibrium action (following any history) under the old lower bound B can be simply scaled up using the ratio of S to B, which is μ. That is, if we replaced the word "sustainable" by the phrase "physical minimum," then the Observation would be trivial. However, S is not a physical minimum. Deviations to asset levels below S are available, and 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 proof of Observation 2 (see the formal statement and proof as Lemma 8 in Section 8) shows that given the concavity of the utility function, such deviations can be suitably deterred. Thus, while S isn't a physical lower bound, it permits us to carry out the same scaling we would achieve if it were.

Let's use Observation 2 to see why the first part of the proposition is true:

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

Recall that $X(A)$ is 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 ever-more-rapid wiggling of $X(A)$ back and forth across the 45^0 line (as $A \downarrow B$) to the more formal arguments — there is $M > B$ such that $X(A) \geq A$ for all $A \in [B, M]$. Figure 3 illustrates this scenario.

Because we are in a non-uniform case, there is 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 $A \in [B, S]$ implies $X(A) \geq A$. Note that *at* S, it must be the case that $X(S)$ *equals* S.[30] Because $X(S) = S$, S is sustainable, though this needs a formal argument.[31]

[30]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 $X(A) \geq A$.

[31]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.

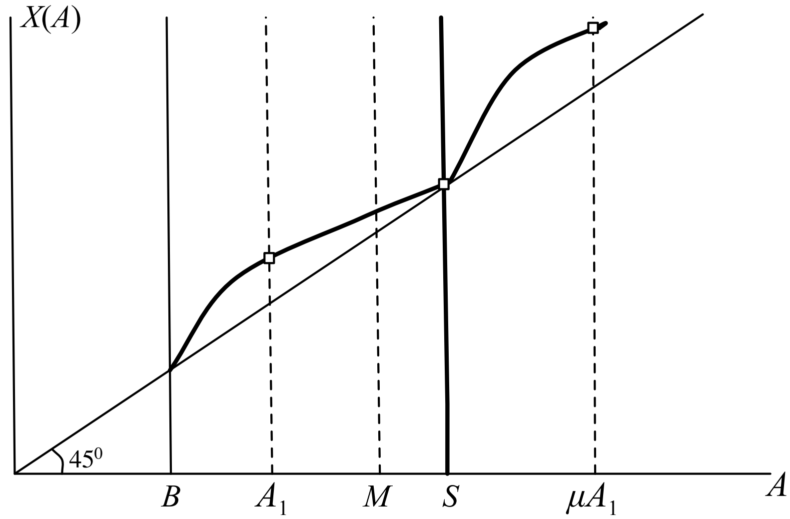


FIGURE 3. ESTABLISHING THE EXISTENCE OF A POVERTY TRAP.

Now Observation 2 implies that $X(A)$ must exceed A just to the right of S: just scale up $X(A_1)$ (for some A_1 close to B) to $\mu X(A_1)$ at μA_1, where $\mu \equiv S/B$. But that is a contradiction to the way we've 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).

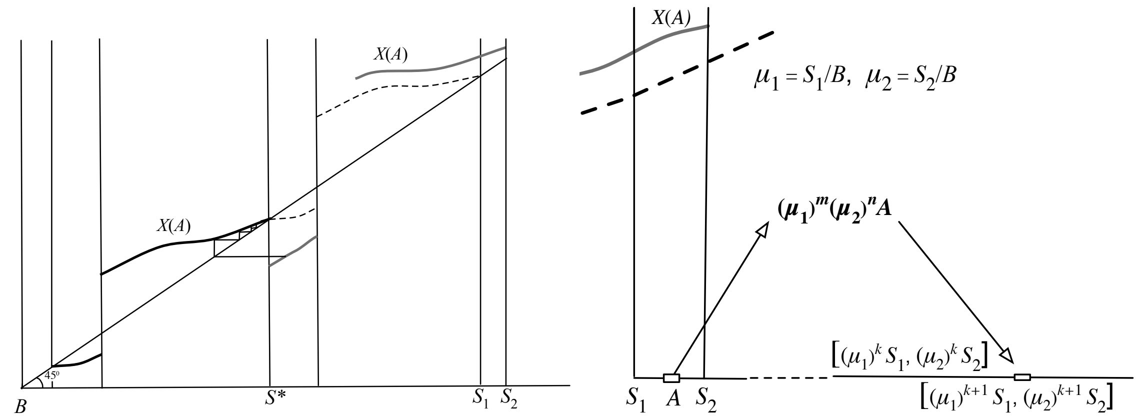
Next, we explain:

(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 (ii), 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 4. Following panel (A) of that figure, begin with a first zone over which $X(A) > A$, and then let S^* be the first asset level thereafter for which $X(A) = A$. As in the exposition for part (i), S^* is sustainable.

By Observation 2, the function $X(A)$ on $[B, S^*]$ can be scaled and replicated as an equilibrium choice over $[S^*, S_1]$, where S_1 bears the same ratio to S^* as S^* does to B.[32]

[32]The actual proof turns considerably more complex at this point. Section 8 makes the complete argument. Briefly, the domain of interest is not exactly $[S^*, S_1]$, but an interval of the form $[S_{**}, S_1]$, where S_{**} *might* coincide with S^* but generally will not. (We proceed here on the assumption that S_{**} does coincide with



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

FIGURE 4. THRESHOLD FOR STRONG SELF-CONTROL.

Figure 4 shows these choices as the dotted line with domain $[S^*, S_1]$. Because there is a poverty trap near B, the line lies below the 45^0 line to the right of B and to the right of S^*. However — and this is at the heart of the argument to be made below — that line does *not* coincide with $X(A)$ on $[S^*, S_1]$.

To see this, consider one feature near S^* that cannot be replicated near B. Just to the right of S^*, one can 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 4 — 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. These lower values do a better job of forestalling deviations at even higher asset levels, and in this way greater punishment ability percolates upward from S*. In particular, for asset levels close to S_1, the incentive constraints are relaxed and larger values of continuation assets (see the solid line segment in this region) are implementable. 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) both S_1 and S_2 are sustainable. Part (a) follows from the fact that $X(S_1) > S_1$ and that X is nondecreasing.

S^*.) There are several associated complications, and the interested reader is referred to Section 8 not just for the formalities, but also for further intuitive discussion.

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 Observation 2, but are actually fully sustainable, given the additional punishment power that has percolated upward from S^*.

Panel B of Figure 4 now focusses on this zone and its implications. The following variation on Observation 2, stated and proved formally as Lemma 16 in Section 8, forms our central argument:

OBSERVATION **3.** *Suppose that 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 observation is illustrated in the second panel of Figure 4. 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 $A' \in (S_1, S_2)$ so that

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

But $X(A') > A'$, so that repeated application of Observation 2 proves that $X(A) > A$. That proves Observation 3.

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

6. SOME ADDITIONAL IMPLICATIONS OF THE THEORY

In this section, we explore the broader implications of our analysis for behavior and policy (aside from the value of "priming the pump" for those caught in the poverty trap). We touch on four topics: first, the effect on saving of easier access to credit; second, the demand for external commitment devices; third, the design of accounts to promote saving; and fourth, the observed variation in marginal propensities to consume from wealth across classes of resource claims.

6.1. **The Effects of Easier Access to Credit.**

It has been widely conjectured that the decline in saving rates among U.S. households during the latter part of 20th century was at least in part attributable to institutional developments that provided progressively easier access to credit.[33] Conventional theory would indeed suggest that more abundant (and cheaper) credit could be expected to reduce aggregate saving. However, in the context of our model, the effects of relaxing credit constraints are more nuanced.

The central theme in this paper is that there is a systematic link between credit limits and the ability to exercise self-control. When an individual's net worth is near the smallest level consistent with credit constraints, he has little scope for disciplining himself through personal rules that "punish" profligacy with decumulation. In contrast, when the same individual has sufficient wealth, it may become feasible for him to adopt and adhere to personal rules that support sustained accumulation.

Within the context of our model, comparative statics with respect to the level of the borrowing constraint (B) are straightforward. Although a fixed borrowing constraint destroys scale neutrality, the model remains scale-neutral in the sense that the ratio of assets A to the constraint B fully determines an individual's ability to exercise self-control. Indeed, we can restate all our observations in terms of this ratio. In particular, Proposition 4 can be interpreted as showing that there are two values, μ' and μ'', with $1 < \mu' \leq \mu'' < \infty$, such that a poverty trap exists whenever $A/B < \mu'$, while unlimited accumulation is possible whenever $A/B > \mu''$.

It follows that the effect on saving of relaxing the credit limit depends on the level of initial assets, A, and is thus ambiguous. The direct effect of such a relaxation is to reduce B, e.g., from B_1 to $B_2 < B_1$, thereby increasing the ratio A/B for each and every individual. That change may allow an individual to escape the poverty trap (i.e., if $A/B_1 < \mu' < A/B_2$), and may even enable him to accumulate assets indefinitely (i.e., if in addition $A/B_2 > \mu''$). However, there is also a downside to easy credit: those whose assets remain below $\mu' B_2$ will slide into an even deeper poverty trap. In any given context, either the first effect or the second may be more prevalent. Notably, Karlan and Zinman (2011) report the results of a field experiment showing that expanded access to costly consumer credit in South Africa on average improved economic self-sufficiency, intra-household control, community status, and overall optimism.

[33]See, e.g., Bacchetta and Gerlach (1997), Ludvigson (1999), Parker (2000) and Glick and Lansing (2011).

6.2. **The Demand for Commitment Devices.** Over the last few decades, time inconsistency has emerged as a central theme in behavioral economics. Yet any consumer sufficiently self-aware to notice her time-inconsistent tendencies should exhibit a demand for precommitment technologies. At a minimum, consumers should acquire such self-awareness with respect to frequently repeated activities for which they consistently fail to follow through on prior intentions. As noted in Section 1, a demand for precommitment has indeed been documented for poor households in developing countries. However, there is surprisingly little evidence that this demand is more widespread,[34] and so nagging doubts about the importance of time inconsistency persist. Skeptics wonder why, if time inconsistency is so prevalent, the market provides few commitment devices, and why unambiguous examples in the field are so difficult to find.

Our analysis provides a potential resolution to this puzzle. Because full-precommitment is neither possible nor desirable (due to the value of flexibility), people must rely to some extent on internal mechanisms for self-control. Significantly, the use of external commitments may undermine the efficacy of those internal mechanisms by rendering effective personal rules infeasible. As an illustration, consider an external commitment that "locks up" assets in an illiquid savings account. The direct effect of that commitment is to increase B, the lower bound on net worth, say from B_1 to $B_2 > B_1$. The impact on saving is then the same as for a tightening of the credit constraint. In particular, defining μ' and μ'' as above, for an individual with $A/B_1 > \mu'' > A/B_2$ the external commitment could impair internal self-control to the point where indefinite accumulation becomes impossible. If in addition $\mu' > A/B_2$, the failure of self-control is so severe that the individual is trapped into depleting assets except those made illiquid via the external commitment mechanism. Accordingly, such individuals have powerful reasons to avoid (partial) external commitments.

In our model, the individuals who value external commitments are those who are asset-poor relative to their credit limits. The asset-rich would rather save on their own. By the same reasoning, if we assume that B is a constant fraction of permanent income, the *income*-rich would exhibit a desire for external commitment, while the income-poor

[34]Studies documenting a demand for precommitment in *developed* countries are scarce. Exceptions include Ariely and Wertenbroch (2002) on homework assignments, Beshears, Choi, Laibson, and Madrian (2011) on commitment savings devices in the U.S., and Houser et al. (2010) for a laboratory experiment in which subjects gain relevant experience. Gine, Karlan, and Zinman (2010) write that "there is little field evidence on the demand for or effectiveness of such commitment devices." For recent surveys, see Bryan et al. (2010) and DellaVigna (2009).

would prefer to rely on internal mechanisms. 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 is, in principle, testable.

6.3. **Designing Accounts to Promote Saving.** Policy makers often try to encourage saving by establishing special accounts for specific purposes, such as retirement, education, medical expenses, or the purchase of a home. Virtually all such accounts entail commitments, but the nature of those commitments differs considerably across programs. As an example, consider retirement savings programs. In almost all cases, funds are to some degree "locked up" until retirement, but the degree of lock-up varies. For public pension programs (e.g., social security) and many private plans (especially of the defined benefit variety), lock-up is absolute. For IRAs it is enforced by a moderate early withdrawal penalty of 10%. For 401(k)s and 403(b)s, the same 10% penalty applies, but employers can also impose additional restrictions and, as an example, often limit such withdrawals to funds contributed by the employee. After retirement, the lock-up continues in a modified form for public pension programs and many private plans: income is paid out at a specified rate, or investment in annuities is mandated. In contrast, IRAs and many other private plans effectively unlock the funds at retirement, making them completely liquid. In addition, participants in retirement savings programs often precommit to contributions. For social security and many private plans, contributions are inflexible. For 401(k)s and 403(b)s, they are adjustable, but only with a significant lag (e.g., a pay period). Only IRA contributions are fully flexible.

Our analysis potentially sheds light on the ways in which savings are affected by the commitment features of special savings accounts. Caution is warranted, inasmuch as the model lacks a retirement period, and therefore maps imperfectly to a realistic life-cycle planning problem. Still, one can interpret it as providing a stylized representation of saving decisions during the accumulation phase of the life cycle.

Following the logic of Section 6.2, it would appear that lock-in has both an upside and a downside. The upside is that it can compensate for the absence of self-control when assets are low; the downside is that it can undermine internal self-control mechanisms when assets are high. Because these effects materialize at different asset levels, it is in principle possible to design programs that capitalize on the upside while avoiding the downside. Intuitively, it would seem that a policy could accomplish that dual objective by requiring the individual to lock up all funds until some asset target is achieved, at which point the lock is removed (irreversibly) and all funds become liquid. Because the poverty trap threshold presumably varies from person to person, each individual would ideally be allowed to select his or her own threshold. Pilot programs with such features have indeed been tested in developing countries.[35]

Formalizing the preceding intuition is less straightforward than one might think. Within the context of our simple model, lock-up would prevent people with low assets from decumulating, but it would not necessarily enable them to employ personal rules that support contributions to the account in the first place, particularly inasmuch as lock-up tends to moderate punishments. Furthermore, there is an obviously superior policy alternative: if we simply allow participants to select (and commit to) their contributions one period in advance, the Ramsey outcome will be achievable from all asset levels.

Despite these issues, our intuition concerning account design is borne out in a slightly more elaborate model that incorporates preference shocks (e.g., reflecting transient needs associated with illnesses requiring costly medical care). In such cases, an exclusive reliance on external commitment is unwarranted and our intuitive statements come more fully into play. Suppose in particular that flow utility is given by

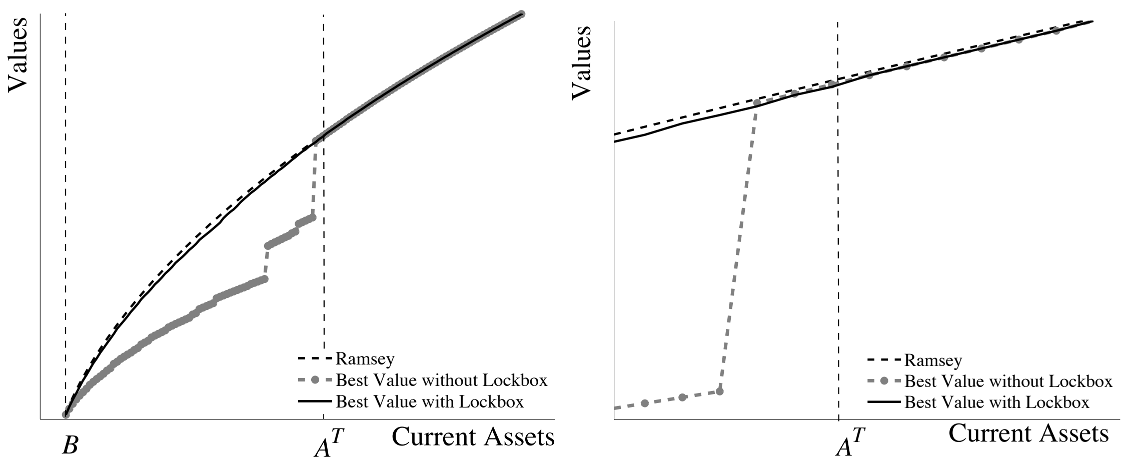
$$u(c, \eta) = \eta \frac{c^{1-\sigma}}{1-\sigma},$$

where η is an iid random variable realized at the outset of each period. In such settings, committing to contributions one period in advance sacrifices the individual's ability to condition consumption on the realization of η, and consequently does not automatically deliver the generalized Ramsey solution. Moreover, if the distribution of η encompasses sufficiently low values, the individual will contribute to a lock-up account in some states of nature even when assets are low.

Due to the complexity of the extended model, we analyze it computationally rather than analytically. For details, see the Appendix.[36] Numerical solutions generally confirm our intuition. Figures 5 and 6 depict results for an illustrative case. Figure 5, panel (A), shows the highest achievable equilibrium value as a function of initial assets for two policy regimes: in the first, only a standard savings account is available; in the second, the individual has access to a lock-up account that unlocks once an appropriately

[35] See Ashraf, Karlan, and Yin (2006), as well as Karlan, McConnell, Mullainathan, and Zinman (2010).

[36] For the parameters, we take $A = 1.3$, $\sigma = 0.5$, $\delta = 0.8$, and $\beta = 0.4$. The taste shock η takes two values, 0.8 (with probability 0.3) and 1.1 (with probability 0.7). The horizontal axis starts at $B = 0.5$, and υ is taken to be a tiny number.



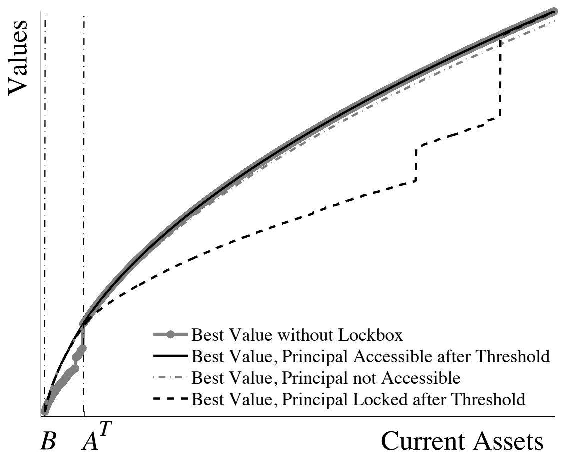
(A) Overall (B) Zoom

FIGURE 5. EQUILIBRIUM VALUES: LOCKBOX WITH UNLOCKING.

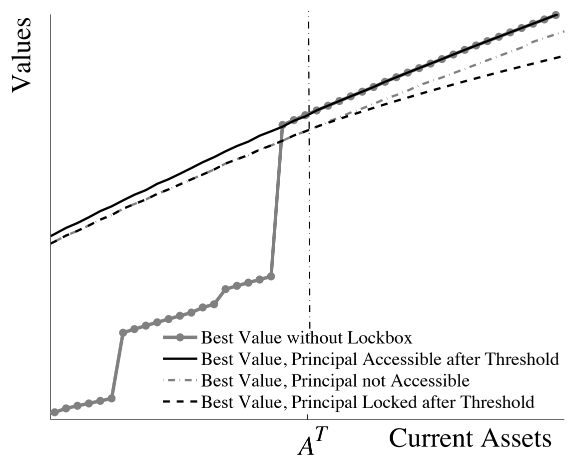
chosen threshold is reached. The "standard model" exhibits a jump in highest value once an individual can effectively save on her own. We've chosen the exogenous lockbox threshold (shown by A^T) so that it is slightly higher than the jump point: anything lower, and the agent will slide back into the poverty trap once the account is unlocked.

At "low" asset levels below the jump point, the individual fares better under the regime with the lock-up account than the one with the standard account. For asset values that exceed the lockbox threshold, there is effectively no lockbox any longer and the two curves must obviously coincide. The figure also shows the value function for the generalized Ramsey solution. Notice that the lock-up regime allows the individual to achieve outcomes close to that theoretical maximum. In our example, it doesn't quite reach that limit, and panel (B) of Figure 5, which amplifies the value functions around the threshold, shows this clearly.

Observe that once the jump point is crossed, the individual can save on her own. Having a threshold that strictly exceeds the jump point creates an interval over which the continuing lock-up may undermine the effectiveness of personal rules, thereby inflicting losses on the individual. Thus, *both* the lock-up and its subsequent release are important. Figure 6 replicates the highest equilibrium value function for the lock-up policy regime, and adds two new lines, representing the highest value attainable under two alternative



(A) Overall



(B) Zoom

FIGURE 6. ALTERNATIVE LOCKBOX REGIMES.

regimes. For the first of these, we eliminate the threshold: principal in the special account remains locked up forever, but the individual can always withdraw current interest. For the second regime, we assume as before that contributions to the lock-up account stop once the threshold is reached (with subsequent saving placed into conventional accounts), but the principal remains locked up forever.

Comparisons with these regimes illustrate the value of both the threshold and the unlocking feature. Values from the first regime are depicted by the dot-dash line in Figure 6; one might view this as the equilibrium values generated by a huge threshold. The figure shows that this regime reduces equilibrium value relative to the "lockbox with unlock" policy of Figure 5 (and panel (B) amplifies the area around the jump point for clarity). A lockbox is needed, but it must be dismantled as well, so as to permit personal rules to come into play.

The second regime illustrates the importance of fully unlocking the lockbox. Here, the individual does have access to conventional savings devices after the threshold, but the principal in the lockbox remains locked. It turns out that the half-step towards unlocking could be even worse (over a subdomain of A) than not unlocking at all; this is shown by the lower dotted line in Figure 6. Briefly, the failure to free the principal is equivalent to a scaling-up of B to the threshold A^T, which creates a new "poverty trap" (relative to the scaled-up B, that is). In short, given the half-measure, our exercise shows that it might be better to keep the lockbox active, and not have a threshold to begin with.[37]

A full analysis of lockbox regimes, with and without unlocking, is beyond the scope of the present paper. But it is hoped that this preliminary analysis will provide a different perspective on the design of such accounts in the presence of hyperbolic agents.

6.4. **Asset-Specific Marginal Propensities to Consume.** A final implication is that the model naturally generates different marginal propensities to consume across classes of resource claims (e.g., between income flows and liquid assets). This phenomenon is documented in Hatsopoulos, Krugman and Poterba (1989), Thaler (1990) and Laibson (1997), though admittedly the empirical evidence for it may be somewhat debatable. To understand the implication, recall from Section 2.1 that we may interpret B, the lower bound on assets, as some function of permanent income, presumably one that is related to the fraction of future labor income that lenders can seize in the event of a default. In other words, 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:

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

while

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

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

In contrast, consider an equivalent jump in y, so that A_t rises by the same amount. Under our specification, B/y is constant so that A/B must fall. According to the ratio interpretation of Proposition 4, self-control is damaged: therefore, the marginal propensity to consume from an unforeseen change in permanent income will be high. Indeed, as long as B is an *increasing* function of permanent income (even if it is more complex), there will be a tendency to observe a higher marginal propensity to consume from permanent

[37]That isn't to say that the half-step is invariably worse than the "fully locked" policy; indeed, that isn't the case even in this example.

income than from liquid assets. Accordingly, our theory provides a new perspective on the "excess sensitivity" of consumption to income.

7. Conclusion

It is evident that if individuals fundamentally differ in their capacities for exercising self-control in intertemporal choices, then the more impulsive of them are likely to end up with poorer asset positions. There is little we have to say about a worldview of poverty that is anchored on the premise of intrinsic differences. What we emphasize, in contrast, is a notion of poverty that feeds back to the capacity for self-control. One way to describe this view is that all individuals have the same *mapping* that runs from their economic position to their behavioral proclivities (in this case, their ability to exercise self-control). The shape of that mapping will determine whether initial poverty (or wealth) subsequently eliminates or amplifies those initial states. In line with a recent and growing literature that emphasizes hysteresis in a variety of settings, we find that poverty damages self-control, while wealth can sustain it. This leads to a new and complementary notion of history-dependence that is rich both in description as well as in its implications for policy.

Specifically, we study a standard model of intertemporal allocation. Agents have quasi-hyperbolic preferences and therefore exhibit present bias (or "impulsiveness"). They seek to control such biases using a system of personal rules (Ainslie 1975, 1992), which we interpret here as history-dependent equilibrium strategies in an intrapersonal dynamic game. Our model is deliberately set up for scale neutrality: the returns to investment are linear, and preferences exhibit constant elasticity. The one feature that breaks this neutrality is an imperfect capital market, modeled as the existence of a strictly lower bound on assets. Our main result is that scale neutrality is broken in favor of the rich and against the poor: there is an asset threshold above which unbounded accumulation is possible, whereas there is a threshold below which the individual must spiral into poverty trap.

Our analysis fits into a large ambient literature on poverty traps, behavioral and otherwise, to which we have referred at various points in this paper. But our particular focus on the links between poverty and self-control deserves further attention along a number of different avenues. Most specifically, our main theorem leaves open the possibility that between the thresholds that define a poverty trap and unbounded accumulation, there may be an intermediate zone which displays neither the inevitability of a poverty trap nor the ability to accumulate in a sustained fashion. Whether that zone is empty or not is an open question which would be important to settle.

Next, while our analysis points to some intriguing relationships between external commitment devices (such as fixed deposit or lockbox retirement accounts) and the efficacy of personal rules, a systematic analytical study of these relationships remains beyond the scope of the present paper. In particular, it would be interesting to study the decision to adopt external commitments for saving, and determine the asset and income levels at which the demand for such devices is maximal.

Finally, and at the broadest level, this paper is a contribution to the behavioral economics of poverty, a subject on which there has been recent empirical focus but little theoretical work. Self-control is one of several behavioral features; others include internally and socially generated aspirations, the reliance on role models, decisions to acquire systematic knowledge about the rate of return from investments in health and education, and informational and psychological distortions that are caused more generally by conditions of poverty. Which of these features amplify initial conditions, and which work to nullify those conditions and create convergence? A taxonomy of behavioral economics along these lines would be of immense significance.

8. Proofs

LEMMA **1.** *For any equilibrium continuation* $\{x, V\}$ *at* A,

$$\begin{aligned} V &\geq \left[u\left(A-\frac{B}{\alpha}\right)+\delta L(B)\right]+\frac{1-\beta}{\alpha\beta}u'\left(A-\frac{B}{\alpha}\right)(x-B) \\ &\geq \left[u\left(A-\frac{B}{\alpha}\right)+\frac{\delta}{1-\delta}u\left(\frac{\alpha-1}{\alpha}B\right)\right]+\frac{1-\beta}{\alpha\beta}u'\left(A-\frac{B}{\alpha}\right)(x-B). \end{aligned}$$

Proof. The payoff associated with $\{x, V\}$ is $(1-\beta)u\left(A-\frac{x}{\alpha}\right)+\beta V$, so

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

because (x, V) is an equilibrium. With u concave, it follows that

$$\begin{aligned} V &\geq \left[u\left(A-\frac{B}{\alpha}\right)+\delta L(B)\right]+\frac{1-\beta}{\beta}\left[u\left(A-\frac{B}{\alpha}\right)-u\left(A-\frac{x}{\alpha}\right)\right] \\ &\geq \left[u\left(A-\frac{B}{\alpha}\right)+\delta L(B)\right]+\frac{1-\beta}{\alpha\beta}u'\left(A-\frac{B}{\alpha}\right)(x-B). \end{aligned} \tag{16}$$

By (5) and $A_t \geq B$ at any date t, we have $u(c_t) \geq u(\upsilon B)$ for any c_t at date t, so that $L(A) \geq (1-\delta)^{-1}u(\upsilon B) > -\infty$. Now, by applying (16) to $A = B$ and $V = L(B)$, or (if needed) a sequence of equilibrium values in $\mathcal{V}(B)$ that converge down to $L(B)$,

$$L(B) \geq u\left(B-\frac{B}{\alpha}\right)+\delta L(B). \tag{17}$$

Combining (16) and (17), the proof is complete. ∎

Proof of Observation 1. This is an immediate consequence of Lemma 1. ∎

Proof of Proposition 1. Claim: if $\mathcal{W}$ is nonempty, has closed graph, and satisfies (8), then it generates $\mathcal{W}'$ with the same properties (plus convex-valuedness). We first prove that $\mathcal{W}'$ is nonempty-valued. Consider the function $H_{\mathcal{W}}$ on $[B, \infty)$ defined by $H_{\mathcal{W}}(A) \equiv \max \mathcal{W}(A)$ for all $A \geq B$. It is easy to see that $H_{\mathcal{W}}$ is usc. It follows that the problem $\max_{x \in [0, \alpha(1-\upsilon)A]} u(A - x/\alpha) + \beta\delta H_{\mathcal{W}}(x)$ is well-defined and admits a (possibly non-unique) solution for every $A \geq B$. Let $x(A)$ denote some solution at A, and define

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

Clearly, w is supported at A by $\mathcal{W}$. (9) is satisfied: pick $x = x(A)$ and $V = H_{\mathcal{W}}(x(A))$. And (10) is satisfied: for each alternative x', take V' to be any element of $\mathcal{W}(x')$.

Claim: $\mathcal{W}'$ has closed graph. Take any sequence $\{A_n, w_n\}$ such that (i) w_n is supported at A_n by $\mathcal{W}$ for all n, and (ii) $(A_n, w_n) \to (A, w)$ (finite) as $n \to \infty$; then w is supported at A by $\mathcal{W}$. To see this, note that for each n, there is x_n feasible for A_n and value $V_n \in \mathcal{W}(x_n)$ such that (9) and (10) are satisfied. Obviously $\{x_n, V_n\}$ is a bounded sequence; pick any limit point (x, V). Then x is certainly a feasible asset choice at A, and $V \in \mathcal{W}(x)$ (because $\mathcal{W}$ has closed graph by assumption). Using the continuation (x, V) at A, it is immediate that (9) is satisfied for w. To prove (10), let x' be any feasible asset choice at A. Then there is $\{x'_n\}$, with x'_n feasible for A_n for all n, such that $x'_n \to x'$. Because w_n is supported at A_n by $\mathcal{W}$, and (x_n, V_n) satisfies (10), there is $V'_n \in \mathcal{W}(x'_n)$ such that

$$u\left(A_n - \frac{x_n}{\alpha}\right) + \beta\delta V_n \geq u\left(A_n - \frac{x'_n}{\alpha}\right) + \beta\delta V'_n \tag{18}$$

for every n. Let V' be any limit point of $\{V'_n\}$. Then, because $\mathcal{W}$ has closed graph, $V' \in \mathcal{W}(x')$. Choose an appropriate subsequence of n such that $\{x'_n, V'_n\}$ converges to (x', V'). Passing to the limit in (18), we must conclude that (10) holds for (A, w) at x'.

These arguments prove that claim that the limit value w is supported at A by $\mathcal{W}$. With the claim in hand, by taking suitable convex combinations it is easy to prove that the correspondence $\mathcal{W}'$ generated by $\mathcal{W}$ has closed graph. It is trivially convex-valued.

Now, consider the sequence $\{\mathcal{V}_k\}$. Because $\mathcal{V}_0$ is nonempty-valued with closed graph, and satisfies (8), the same is true of the $\mathcal{V}_k$'s. Moreover, for each $t \geq 0$ and all $A \geq B$,

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

Take infinite intersections of these nested compact sets (at each A) to argue that

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

is nonempty for every A. Furthermore, because $\mathcal{V}_k(A)$ is convex for all $k \geq 0$, so is $\mathcal{V}^*(A)$. Moreover, $\mathcal{V}^*$ has compact graph on any compact interval $[B, D]$,[38] and therefore it has closed graph everywhere. We will show that $\mathcal{V}^*$ generates itself. To this end, we

[38]On any compact interval, the (restricted) graphs of the $\mathcal{V}_k$'s are compact and their infinite intersection is the graph of $\mathcal{V}^*$ on the same interval, which must then be compact.

first show for each A, every w supported at A by $\mathcal{V}^*$ lies in $\mathcal{V}^*(A)$. Pick such a value w at A. Then there is a feasible continuation asset choice x at A and $V \in \mathcal{V}^*(x)$ such that (9) holds, and for every feasible choice x' at A, there is $V' \in \mathcal{V}^*(x')$ such that (10) holds. But these continuations are available in $\mathcal{V}_k$ for every k, which means that w is supported at A by every $\mathcal{V}_k$. It follows that $w \in \mathcal{V}_{k+1}(A)$ for every k, so that $w \in \mathcal{V}^*(A)$.

We complete the argument by showing that for every A, $\max \mathcal{V}^*(A)$ and $\min \mathcal{V}^*(A)$ are supportable at A by $\mathcal{V}^*$.[39] The same argument works in either case, so we show this for $\max \mathcal{V}^*(A)$. Because $\mathcal{V}^*(A) = \bigcap_{t=0}^{\infty} \mathcal{V}_k(A)$, the sequence of values $w_k \equiv \max \mathcal{V}_k(A)$ converges to $H(A)$. Moreover, w_k cannot be a proper convex combination of other values in $\mathcal{V}_k(A)$, so w_k is supportable at A by $\mathcal{V}_k$, for every k. That is, for each k, there is x_k feasible for A and value $V_k \in \mathcal{V}_k(x_k)$ such that (9) and (10) are satisfied for w_k. It is easy to see that $\{x_k, V_k\}$ is a bounded sequence. Pick any limit point (x, V) of $\{x_k, V_k\}$. Then x is a feasible choice at A, and $V \in \mathcal{V}^*(x)$.[40] Using the continuation (x, V) at A, then, (9) is satisfied for $w = \max \mathcal{V}^*(A)$ (under $\mathcal{V}^*$).

Now, let x' be any feasible asset choice at A. Because w_k is supported at A by $\mathcal{V}_k$, and (x_k, V_k) has been chosen such that (10) is satisfied, there exists $V'_k \in \mathcal{V}_k(x')$ such that

$$u\left(A - \frac{x_k}{\alpha}\right) + \beta\delta V_k \geq u\left(A - \frac{x'}{\alpha}\right) + \beta\delta V'_k \tag{19}$$

for every k. Let V' be any limit point of $\{V'_k\}$. Then, by the argument already used (see footnote 40), $V' \in \mathcal{V}^*(x')$. Choose an appropriate subsequence of n such that $\{x'_n, V'_n\}$ converges to (x', V'). Passing to the limit in (19), we see that (10) holds for (A, w) at x'.

This shows that $\mathcal{V}^*$ generates $\mathcal{V}^*$. It is immediate that $\mathcal{V}^*$ contains every correspondence that generates itself,[41] so it is the same as our equilibrium correspondence $\mathcal{V}$. ■

Given Proposition 1, let $H(A)$ and $L(A)$ be the maximum and minimum values of the equilibrium value correspondence $\mathcal{V}$. Because the graph of $\mathcal{V}$ is closed, H is usc and L is lsc. In what follows we take care to account for possible discontinuities in L, which are

[39] Because $\mathcal{V}^*(A)$ is convex, it equals $[\min \mathcal{V}^*(A), \max \mathcal{V}^*(A)]$. We've shown that all w supportable at A by $\mathcal{V}^*$ must indeed lie in $\mathcal{V}^*(A)$. So, provided we can show that $\max \mathcal{V}^*(A)$ and $\min \mathcal{V}^*(A)$ are supportable at A by $\mathcal{V}^*$, it must follow that $\mathcal{V}^*(A)$ is the convex hull of *all* values supported at A by $\mathcal{V}^*$.

[40] To see why, pick any n in the sequence. Then for $k \geq n$, $V_k \in \mathcal{V}_k(x_k) \subseteq V_n(x_k)$, so that $V \in \mathcal{V}_n(x)$ by the closed-graph property of $\mathcal{V}_n$. It follows that $V \in \mathcal{V}_n(x)$ for *every* n, so that $V \in \mathcal{V}^*(x)$ as asserted.

[41] Let $\mathcal{V}'$ be any self-generating correspondence. Then if $\mathcal{V}' \subseteq \mathcal{V}_k$, we have $\mathcal{V}' \subseteq \mathcal{V}_{k+1}$. But $\mathcal{V}' \subseteq \mathcal{V}_0$, so it follows that $\mathcal{V}' \subseteq \mathcal{V}_k$ for every k, which implies $\mathcal{V}' \subseteq \mathcal{V}^*$.

unfortunately endemic. Let x be a feasible choice of continuation asset at A. Consider all limits of sequences of the form $\{L(x^n)\}$, where $x^n \in [B, \alpha(1-\upsilon)A]$ for all n and $x^n \to x$. Each limit is an equilibrium value at x, because $\mathcal{V}$ has closed graph. Moreover, the collection of all such limits at x (given A) is compact, so a *largest* value $M(x, A)$ exists. That defines the function $M(x, A)$ for $A \geq B$ and $x \in [B, \alpha(1-\upsilon)A]$. An individual can guarantee herself a continuation value that is arbitrarily close to $M(x, A)$, starting from A (by making an asset choice arbitrarily close to x).

LEMMA **2.** *For given A, $M(x, A)$ is usc in x, and for given x, it is nondecreasing in A, and independent of A as long as $x < \alpha(1-\upsilon)A$.*

Proof. Pick x^n feasible for A such that $x^n \to x \in [B, \alpha(1-\upsilon)A]$ and a corresponding sequence $M^n = M(x^n, A)$. Suppose without loss of generality that $M^n \to M$. For each n, there is $y^n \in [B, \alpha(1-\upsilon)A]$ such that $|y^n - x^n| < 1/n$, and $|L(y^n) - M^n| < 1/n$. It is then easy to see that $y^n \to x$ and $L(y^n) \to M$. So M is a limit value at x, which implies $M(x, A) \geq M$. Therefore $M(x, A)$ is usc in x. To prove that $M(x, A)$ is nondecreasing in A, observe that every sequence of the form $\{L(x^n)\}$, where $x^n \in [B, \alpha(1-\upsilon)A]$, is fully available at $A' > A$, whenever it is available at A. It is also obvious that for any x, exactly the same limit values of $\{L(x^n)\}$ are available when $x < \alpha(1-\upsilon)A$, so that $M(x, A)$ is then unchanging in A whenever the strict inequality holds. ∎

Lemma 2 implies that the following "best deviation payoff" at A is well-defined:

$$D(A) = \max_x u\left(A - \frac{x}{\alpha}\right) + \beta\delta M(x, A), \tag{20}$$

where it is understood that $x \in [B, \alpha(1-\upsilon)A]$. Lemma 2 also implies that $D(A)$ is an increasing function. Note that D does not necessarily use worst punishments everywhere, but nonetheless a deviant can get payoff arbitrarily close to $D(A)$. That implies:

LEMMA **3.** *The pair (x, V) is an equilibrium continuation at A if and only if $x \in [B, \alpha(1-\upsilon)A]$, $V \in \mathcal{V}(x)$ and*

$$u\left(A - \frac{x}{\alpha}\right) + \beta\delta V \geq D(A). \tag{21}$$

Proof. Sufficiency: if (x, V) is not an equilibrium continuation, then there exists $y \neq x$ such that $u(A - x/\alpha) + \beta\delta V < u(A - y/\alpha) + \beta\delta L(y)$. But $L(y) \leq M(y, A)$, so $u(A - x/\alpha) + \beta\delta V < u(A - y/\alpha) + \beta\delta M(y, A) \leq D(A)$.

Necessity: if (x, V) is an equilibrium continuation at A, then $x \in [B, \alpha(1-\upsilon)A]$ and $V \in \mathcal{V}(x)$. Moreover, for every feasible y, and sequence of feasible $\{y^n\}$ with $y^n \to y$,

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

where the inequality holds trivially for $y^n = x$ (because $V \geq L(x)$) and by incentive compatibility for $y^n \neq x$. Passing to the limit in that inequality, we must conclude that

$$u\left(A - \frac{x}{\alpha}\right) + \beta\delta V \geq u\left(A - \frac{y}{\alpha}\right) + \beta\delta M(y, A).$$

Maximizing the right hand side of this inequality over y, we obtain the desired result. ■

LEMMA **4.** *If d solves* (20)*, then $\{d, M(d, A)\}$ is an equilibrium continuation at A.*

Proof. Because $\mathcal{V}$ has closed graph, $M(d, A) \in \mathcal{V}(d)$. Now apply Lemma 3. ■

LEMMA **5.** *$L(A)$ is increasing on $[B, \infty)$.*

Proof. Let $A'' > A' \geq B$. Consider the equilibrium that generates value $L(A'')$ starting from A'', with associated continuation $\{A_1'', V''\}$. By Lemma 3,

$$u\left(A'' - \frac{A_1''}{\alpha}\right) + \beta\delta V'' \geq u\left(A'' - \frac{x}{\alpha}\right) + \beta\delta M(x, A'') \tag{22}$$

for $x \in [B, \alpha(1-\upsilon)A'']$. It follows that $V'' > M(x, A'')$ for all $x < A_1''$, which implies

$$L(A'') = u\left(A'' - \frac{A_1''}{\alpha}\right) + \delta V'' > u\left(A'' - \frac{x}{\alpha}\right) + \delta M(x, A'') \tag{23}$$

for all $x < A_1''$. Now construct an equilibrium from A': the choice A_1'' (if feasible) is followed by V'', while each other $x \in [B, \alpha(1-\upsilon)A']$ is followed by $M(x, A')$.[42] Note that

$$\begin{aligned} u\left(A' - \frac{A_1''}{\alpha}\right) + \beta\delta V'' \quad &> \quad u\left(A' - \frac{x}{\alpha}\right) + \beta\delta M(x, A'') \\ &\geq \quad u\left(A' - \frac{x}{\alpha}\right) + \beta\delta M(x, A'), \end{aligned} \tag{24}$$

for $x \in (A_1'', \alpha(1-\upsilon)A']$ (assuming this set is non-empty), where the first inequality uses the strict concavity of u, $A' < A''$ and (22), and the second uses Lemma 2.

To complete the description of equilibrium, we must choose a particular continuation at A': pick continuation $\{y, V\}$ to maximize payoff over the specified continuations above.

[42] Recall that $M(x, A')$ is indeed an equilibrium value at x because $\mathcal{V}$ has closed graph.

Given (24), that is tantamount to choosing from the greatest of the payoffs

$$u\left(A' - \frac{x}{\alpha}\right) + \beta\delta M(x, A')$$

for $x \in [B, \min\{\alpha(1-\upsilon)A', A_1''\}]$, and the payoff at $x = A_1''$ (if feasible), which is

$$u\left(A' - \frac{A_1''}{\alpha}\right) + \beta\delta V'',$$

and a solution is well-defined, because M is usc in x, and the replacement of $M(A_1'', A)$ by V'' at A_1'' (if feasible for A') only increases payoff. The chosen continuation $\{y, V\}$ must be an equilibrium, and by (24), $y \leq A_1''$. If $y < A_1''$, then by (23) and Lemma 2,

$$L(A'') > u\left(A'' - \frac{y}{\alpha}\right) + \delta M(y, A'') > u\left(A' - \frac{y}{\alpha}\right) + \delta M(y, A') \geq L(A'),$$

and if $y = A_1''$, then again

$$L(A'') = u\left(A'' - \frac{A_1''}{\alpha}\right) + V'' > u\left(A' - \frac{y}{\alpha}\right) + V'' \geq L(A').$$

So in both cases, $L(A'') > L(A')$, as desired. ■

Lemma 5 makes it easy to visualize $M(x, A)$. With L increasing, let $L^+(A)$ denote the right hand limit of L at A; i.e., the common limit of all sequences $\{L(A^n)\}$ as $A^n \downarrow A$, with $A^n > A$ for all n. Clearly, L^+ is an increasing, right-continuous function.

LEMMA **6.** *For any A and $x \in [B, \alpha(1-\upsilon)A)$, $M(x, A)$ equals $L^+(x)$. At $x = \alpha(1-\upsilon)A$, it equals $L(x)$.*

Proof. Obvious, given Lemma 5 and the definitions of L and M. ■

LEMMA **7.** (a) *Let $d(A)$ solve* (20). *If $A_1 < A_2$, then $d(A_1) \leq d(A_2)$. Moreover, a largest solution $d^*(A)$ is well-defined for each A, and it is nondecreasing in A.*

(b) *$d^*(A)$ is right-continuous at any A such that* $\lim_n d^*(A^n) < \alpha(1-\upsilon)A$ *for $A^n \downarrow A$.*

Proof. Let $x_i \equiv d(A_i)$ for $i = 1, 2$. Suppose, on the contrary, that $x_1 > x_2$. Notice that x_1 is feasible at A_2 (because $A_1 < A_2$ and x_1 is feasible at A_1), and that x_2 is feasible at A_1 (because $x_2 < x_1$). Therefore

$$u\left(A_i - \frac{x_i}{\alpha}\right) + \beta\delta M(x_i, A_i) \geq u\left(A_i - \frac{x_j}{\alpha}\right) + \beta\delta M(x_j, A_i)$$

for $i = 1, 2$ and $j \neq i$. Combining these two inequalities, and using Lemma 2 to conclude that $M(x_1, A_2) \geq M(x_1, A_1)$, while $M(x_2, A_2) = M(x_2, A_1)$,[43]

$$\left[u\left(A_2 - \frac{x_2}{\alpha}\right) - u\left(A_2 - \frac{x_1}{\alpha}\right)\right] \geq \left[u\left(A_1 - \frac{x_2}{\alpha}\right) - u\left(A_1 - \frac{x_1}{\alpha}\right)\right].$$

But the above inequality contradicts the strict concavity of u. So $x_1 \leq x_2$, as desired.

Next we show that a largest maximizer $d^*(A)$ exists at each A. Let d^n each solve (20) at A, and say that $d^n \to d$. Because $M(x, A)$ is usc in x (Lemma 2),

$$\lim_{n\to\infty} u\left(A - \frac{d^n}{\alpha}\right) + \beta\delta M(d^n, A) \leq u\left(A - \frac{d}{\alpha}\right) + \beta\delta M(d, A),$$

but the left-hand side of this inequality *is* the maximized value of (20) for every n, so the right-hand side must have the same value, which shows that d also solves (20). That proves the existence of a largest maximizer $d^*(A)$ at every A, and the arguments so far show that $d^*(A)$ is nondecreasing, so the proof of part (a) is complete.

For part (b), fix A and let $d \equiv \lim_n d^*(A^n) < \alpha(1 - \upsilon)A$ for $A^n \downarrow A$ (noting that $\{d^*(A^n)\}$ is monotone). Clearly, d is feasible at A. To prove the right continuity of d^* at A, we show that d maximizes (20) at A. Suppose not. Let d' maximize (20) at A; then

$$u\left(A - \frac{d'}{\alpha}\right) + \beta\delta M(d', A) > u\left(A - \frac{d}{\alpha}\right) + \beta\delta M(d, A). \tag{25}$$

Notice that $d' \leq d$ (by part (a), already proved), so $d' < \alpha(1-\upsilon)A \leq \alpha(1-\upsilon)A^n$ for all n. So by Lemma 2, $M(x, A)$ is independent of A at (d', A), and an analogous assertion is true of A^n. Therefore, not only is d' feasible for all A^n, we also have

$$\lim_n u\left(A^n - \frac{d'}{\alpha}\right) + \beta\delta M(d', A^n) = u\left(A - \frac{d'}{\alpha}\right) + \beta\delta M(d', A). \tag{26}$$

Define $d^n \equiv d^*(A^n)$, and note that for n large, $d^n < \alpha(1 - \upsilon)A \leq \alpha(1 - \upsilon)A^n$. Using the independence of M in A^n and recalling that $M(x, A)$ is usc in x (Lemma 2),

$$\lim_n u\left(A^n - \frac{d^n}{\alpha}\right) + \beta\delta M(d^n, A^n) \leq u\left(A - \frac{d}{\alpha}\right) + \beta\delta M(d, A). \tag{27}$$

Combining (25)–(27), we must conclude that for n large,

$$u\left(A^n - \frac{d'}{\alpha}\right) + \beta\delta M(d', A^n) > u\left(A^n - \frac{d^n}{\alpha}\right) + \beta\delta M(d^n, A^n),$$

which contradicts the fact that d^n maximizes (20) for all n. ∎

[43]Note that $x_2 < x_1 \leq \alpha(1 - \upsilon)A_i$ for $i = 1, 2$. By Lemma 2, $M(x_2, A_2) = M(x_2, A_1)$.

Define the *maintenance value* of an asset level A by $V^s(A) \equiv \frac{1}{1-\delta} u\left(\frac{\alpha-1}{\alpha}A\right)$, and the *maintenance payoff* by $P^s(A) \equiv \left[1+\frac{\beta\delta}{1-\delta}\right] u\left(\frac{\alpha-1}{\alpha}A\right)$. Say that an asset level S is *sustainable* if there is a stationary equilibrium path from S, or equivalently (by Lemma 3) if $P^s(S) \geq D(S)$.

LEMMA **8** (Observation 2 in main text)**.** *Let $S > B$ be a sustainable asset level, and $\mu \equiv S/B$. Then, if $\{A_t^*\}$ is an equilibrium path from A_0:*

(a) *$\{\mu A_t^*\}$ is an equilibrium path from μA_0.*

(b) *For all t with $\mu A_t^* > S$ and for every $A < S$,*

$$u\left(\mu A_t^* - \frac{\mu A_{t+1}^*}{\alpha}\right) + \beta \sum_{s=t+1}^{\infty} \delta^{s-t} u\left(\mu A_s^* - \frac{\mu A_{s+1}^*}{\alpha}\right) > u\left(\mu A_t^* - \frac{A}{\alpha}\right) + \beta\delta M(A, A_t^*).$$

Proof. Part (a). Let policy ϕ sustain $\{A_t^*\}$ from A_0. Define a new policy ψ:

(i) For any $h_t = (A_0 \ldots A_t)$ with $A_s \geq S$ for $s = 0, \ldots, t$, let $\psi(h_t) = \mu\phi\left(\frac{h_t}{\mu}\right)$.

(ii) For h_t with $A_k < S$ for some smallest $k \leq t$, define $h'_{t-k} = (A_k \ldots A_t)$. Let $\psi(h_t) = \phi_\ell(h'_{t-k})$, where ϕ_ℓ is the equilibrium policy with value $L(A_k)$ at A_k.

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

$$P_\psi(h_t) = \mu^{1-\sigma} P_\phi\left(\frac{h_t}{\mu}\right) \text{ and } V_\psi(h_t) = \mu^{1-\sigma} V_\phi\left(\frac{h_t}{\mu}\right). \tag{28}$$

We now show that ψ is an equilibrium. First, consider any h_t such that $A_k < S$ at some first $k \leq t$. Then as of period k the policy function ψ shifts to the equilibrium with 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

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

Multiplying through by $\mu^{1-\sigma}$ and using (28), we see that

$$P_\psi(h_t) \geq u\left(A_t - \frac{A}{\alpha}\right) + \beta\delta V_\psi(h_t.A), \tag{29}$$

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

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

$$P^s(S) \geq D(S) \geq u\left(S - \frac{A}{\alpha}\right) + \beta\delta M(A, S) \tag{30}$$

by Lemma 3. At the same time, (29) applied to $A = S$ implies

$$P_\psi(h_t) \geq u\left(A_t - \frac{S}{\alpha}\right) + \beta\delta V_\psi(h_t.S). \tag{31}$$

Using (28) along with $L(B) \geq V^s(B)$ (see Observation 1), (31) becomes

$$\begin{aligned} P_\psi(h_t) &\geq u\left(A_t - \frac{S}{\alpha}\right) + \beta\delta\mu^{1-\sigma} V_\phi\left(\frac{h_t}{\mu}.B\right) \\ &\geq u\left(A_t - \frac{S}{\alpha}\right) + \beta\delta\mu^{1-\sigma} L(B) \\ &\geq u\left(A_t - \frac{S}{\alpha}\right) + \beta\delta\mu^{1-\sigma} V^s(B) \\ &= u\left(A_t - \frac{S}{\alpha}\right) + \beta\delta V^s(S) \\ &= \left[u\left(A_t - \frac{S}{\alpha}\right) - u\left(S\left(1 - \frac{1}{\alpha}\right)\right)\right] + P^s(S). \end{aligned} \tag{32}$$

Combining (30) and (32),

$$\begin{aligned} P_\psi(h_t) &\geq \left[u\left(A_t - \frac{S}{\alpha}\right) - u\left(S\left(1 - \frac{1}{\alpha}\right)\right)\right] + u\left(S - \frac{A}{\alpha}\right) + \beta\delta M(A, S) \\ &= \left[u\left(A_t - \frac{S}{\alpha}\right) - u\left(S - \frac{S}{\alpha}\right)\right] - \left[u\left(A_t - \frac{A}{\alpha}\right) - u\left(S - \frac{A}{\alpha}\right)\right] \\ &\quad + u\left(A_t - \frac{A}{\alpha}\right) + \beta\delta M(A, S) \\ &\geq u\left(A_t - \frac{A}{\alpha}\right) + \beta\delta M(A, S) \end{aligned} \tag{33}$$

where the second inequality follows from the concavity of u and the fact that $A < S \leq A_t$. But, because $M(A, S) \geq L(A) = V_\psi(h_t.A)$, the right hand side of (33) is at least as large as the payoff from the deviation, which is $u\left(A_t - [A/\alpha]\right) + \beta\delta V_\psi(h_t.A)$. It follows that the deviation A is unprofitable, so that ψ is an equilibrium.

Part (b). The second inequality in (33) holds strictly when $A_t > S$ and $A < S$, by the strict concavity of u. Apply (33) (with strict inequality) at date t, with h_t equal to the history on the equilibrium path and setting $M(A,S) = M(A, A_t^*)$ (Lemma 2). ■

LEMMA **9.** *For any asset level A and any path $\{A_t\}$ with $A_t \leq A$ for all $t \geq 0$,*

$$V^s(A) - \sum_{t=0}^{\infty} \delta^t u\left(A_t - \frac{A_{t+1}}{\alpha}\right) \geq u'\left(\frac{\alpha - 1}{\alpha} A\right)\left(\delta - \frac{1}{\alpha}\right)(A - A_1) \geq 0. \tag{34}$$

Proof. Let Δ stand for the left hand side of (34); then

$$\begin{aligned}
\Delta &= \sum_{t=0}^{\infty} \delta^t \left[u\left(\frac{\alpha-1}{\alpha}A\right) - u\left(A_t - \frac{A_{t+1}}{\alpha}\right)\right] \\
&\geq u'\left(\frac{\alpha-1}{\alpha}A\right) \sum_{t=0}^{\infty} \delta^t \left[A - \frac{A}{\alpha} - A_t + \frac{A_{t+1}}{\alpha}\right] \\
&= u'\left(\frac{\alpha-1}{\alpha}A\right) \sum_{t=0}^{\infty} \delta^t \left[(A - A_t) - \frac{A - A_{t+1}}{\alpha}\right] \\
&= u'\left(\frac{\alpha-1}{\alpha}A\right) \left[(A - A_0) + \left(\delta - \frac{1}{\alpha}\right) \sum_{t=0}^{\infty} \delta^t (A - A_{t+1})\right] \\
&\geq u'\left(\frac{\alpha-1}{\alpha}A\right)\left(\delta - \frac{1}{\alpha}\right)(A - A_1) \geq 0,
\end{aligned}$$

where the first inequality uses the concavity of u and the last uses $\delta\alpha > 1$. ■

Let $X(A)$ be the largest and $Y(A)$ the smallest equilibrium asset choice at A.

LEMMA **10.** *$X(A)$ and $Y(A)$ are well-defined and non-decreasing, and X is usc.*

Proof. By Lemma 3, $X(A)$ (resp. $Y(A)$) is the largest (resp. smallest) value of $A' \in [B, \alpha(1-\upsilon)A]$ satisfying

$$u\left(A - \frac{A'}{\alpha}\right) + \beta\delta H(A') \geq D(A) \tag{35}$$

$X(A)$ and $Y(A)$ are well-defined because H is usc.

To show that X is nondecreasing, pick $A_1 < A_2$. (35) implies that

$$u\left(A_1 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) \geq u\left(A_1 - \frac{y}{\alpha}\right) + \beta\delta L(y)$$

for all $y \in [B, \alpha(1-\upsilon)A]$. It follows from the concavity of u that

$$u\left(A_2 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) \geq u\left(A_2 - \frac{y}{\alpha}\right) + \beta\delta L(y) \tag{36}$$

for all $y \in [B, X(A_1)]$. If the inequality extends to all $y \in [B, \alpha(1-\upsilon)A]$, the claim would be established. Otherwise there is $x' \in (X(A_1), \alpha(1-\upsilon)A_2]$ such that

$$u\left(A_2 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) < u\left(A_2 - \frac{x'}{\alpha}\right) + \beta\delta L(x'). \tag{37}$$

Combine (36) and (37) to see that

$$\begin{aligned} u\left(A_2 - \frac{x'}{\alpha}\right) + \beta\delta L(x') &> u\left(A_2 - \frac{X(A_1)}{\alpha}\right) + \beta\delta H(X(A_1)) \\ &\geq u\left(A_2 - \frac{y}{\alpha}\right) + \beta\delta L(y) \end{aligned} \tag{38}$$

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 $[X(A_1), \alpha(1-\upsilon)A_2]$; in light of (38) 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 $X(A_2) \geq z^* \geq X(A_1)$.

To show that $Y(A)$ is non-decreasing, pick $A_1 < A_2$. If $Y(A_2) \geq \alpha[1-\upsilon]A_1$, we're done, so suppose that $Y(A_2) < \alpha[1-\upsilon]A_1$. Construct an equilibrium from A_1 as follows. For any $A \in [B, Y(A_2)]$, assign the continuation value $H(A)$, and for $A \in (Y(A_2), \alpha[1-\upsilon]A_1]$, assign the continuation value $L(A)$. Finally, for the equilibrium asset choice at A_1, assign A', where A' solves

$$\max_{A \in [B, Y(A_2)]} u\left(A_1 - \frac{A}{\alpha}\right) + \beta\delta H(A)$$

(Because H is usc, a solution exists.) We claim that A' maximizes payoff over all the above specifications, so that $\{A', H(A')\}$ is an equilibrium continuation. It certainly does so over choices in $[B, Y(A_2)]$, by construction. For $A \in (Y(A_2), \alpha[1-\upsilon]A_1]$,

$$u\left(A_2 - \frac{Y(A_2)}{\alpha}\right) + \beta\delta H(Y(A_2)) \geq u\left(A_2 - \frac{A}{\alpha}\right) + \beta\delta M(A, A_2),$$

so by the concavity of u and Lemma 2,

$$\begin{aligned} u\left(A_1 - \frac{Y(A_2)}{\alpha}\right) + \beta\delta H(Y(A_2)) &\geq u\left(A_1 - \frac{A}{\alpha}\right) + \beta\delta M(A, A_2) \\ &\geq u\left(A_1 - \frac{A}{\alpha}\right) + \beta\delta M(A, A_1), \end{aligned}$$

which proves the claim. Because $A' \leq Y(A_2)$, it follows that $Y(A_1) \leq Y(A_2)$.

Finally, we show that X is usc. For any $A^* \geq B$, $\lim_{A \uparrow A^*} X(A) \leq X(A^*)$ because $X(A)$ is nondecreasing. Now consider any decreasing sequence $A^k \downarrow A^*$, and let X^* denote the (well-defined) limit of $X(A^k)$. For each k, $u\left(A^k - X(A^k)/\alpha\right) + \beta\delta H(X(A^k)) \geq D(A^k)$. Because H is usc and $D(A)$ is nondecreasing, $u\left(A^* - X^*/a\right) + \beta\delta H(X^*) \geq \lim_{k\to\infty} D(A^k) \geq D(A^*)$. That implies $X(A^*) \geq X^* = \lim_{A\downarrow A^*} X(A)$. (In fact, because $X(A)$ is non-decreasing, $X(A^*) = \lim_{A\downarrow A^*} X(A)$.) ■

LEMMA **11.** *If $X(A) = A$, then A is sustainable.*

Proof. Let $A_1 = A$ along with some value V_1 be an equilibrium continuation at A. Then

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

by Lemma 3. By Lemmas 9 and 10, $V_1 \leq (1-\delta)^{-1} u\left(\frac{\alpha-1}{\alpha}A\right)$. Using this in the inequality above, we see that $P^s(A) \geq D(A)$, so that A is sustainable. ■

LEMMA **12.** *In the nonuniform case, $\beta\delta(\alpha-1)/(1-\delta) < 1$.*

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

To this end, assume that all "future selves" employ the policy function $\phi(A) = kA$ with $k \in [1, \alpha]$ for all $A \geq B$. 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

$$Q \equiv \frac{(\alpha - k)^{1-\sigma}}{\alpha^{1-\sigma}(1 - \delta k^{1-\sigma})} \tag{39}$$

The corresponding necessary and sufficient first-order condition is

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

After some manipulation, we obtain

$$\frac{A}{x} = \frac{1}{\alpha} + \left(\frac{1}{\alpha\beta\delta Q}\right)^{1/\sigma} \equiv \frac{1}{k^*} \tag{40}$$

Note that $x = k^* A$. Accordingly, the policy function is an equilibrium if $k^* = k$. Substituting (39) into (40) and rearranging yields

$$k^\sigma = \alpha\beta\delta + (1 - \beta)\,\delta k \tag{41}$$

Define $\Lambda(k) \equiv k^\sigma$ and $\Phi(k) = \alpha\beta\delta + (1 - \beta)\,\delta k$. Notice that $\Lambda(1) \leq \Phi(1)$ (given that $\beta\delta(\alpha - 1)/(1 - \delta) \geq 1$), 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. ∎

LEMMA **13.** *Under nonuniformity, the problem*

$$\max_{x \in [0, \alpha(1-\upsilon)A]} \left[u\left(A - \frac{x}{\alpha}\right) + \beta\delta V^s(x)\right].$$

has a unique solution $x(A)$ with $x(A) = \Gamma A$, where $0 < \Gamma < 1$. Moreover, the maximand is strictly decreasing in x for all $x \geq x(A)$.

Proof. The maximand is a continuous, strictly concave function, so it has a unique, continuous solution $x(A)$ for each A. Moreover, by strict concavity, the maximand must strictly decline in x for all $x \geq x(A)$. Define $\xi = \beta\delta(\alpha - 1)/(1 - \delta)$. By nonuniformity and Lemma 12, we have $\xi < 1$. Routine computation reveals that $x(A) = \Gamma A$, where

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

which (given $\sigma > 0$ and $\xi < 1$) implies $\Gamma < 1$. ∎

LEMMA **14.** *For any $A_0 \geq B$, maximize $\sum_{t=0}^{\infty} \delta^t u\left(A_t - \frac{A_{t+1}}{\alpha}\right)$, subject to $A_{t+1} \in [B, \alpha(1 - \upsilon)A_t]$, and $A_{t+1} \leq X(A_t)$ for all $t \geq 0$. Then a solution exists, and any solution path $\{A_t^*\}$ is also an equilibrium path starting from A_0.*

Proof. u is continuous and $X(A_t)$ is usc (Lemma 10), so a solution $\{A_t^*\}$ exists. Let $\{V_t^*\}$ be the sequence of continuation values associated with $\{A_t^*\}$. Consider an equilibrium path from date t, call it $\{A_\tau\}$, sustaining $X(A_t^*)$ at A_t^* and providing continuation value $H(X(A_t^*))$ thereafter. This path necessarily satisfies $A_{\tau+1} \leq X(A_\tau)$ for all $\tau \geq t$, so the definitions of $\{A_t^*\}$ and $\{V_t^*\}$ imply that

$$u\left(A_t^* - \frac{A_{t+1}^*}{\alpha}\right) + \delta V_{t+1}^* \geq u\left(A_t^* - \frac{X(A_t^*)}{\alpha}\right) + \delta H(X(A_t^*)) \tag{42}$$

Also, because $A_{t+1}^* \leq X(A_t^*)$ and $\beta < 1$, we have

$$\left(\frac{1}{\beta} - 1\right) u\left(A_t^* - \frac{A_{t+1}^*}{\alpha}\right) \geq \left(\frac{1}{\beta} - 1\right) u\left(A_t^* - \frac{X(A_t^*)}{\alpha}\right) \tag{43}$$

Adding (42) to (43) and multiplying through by β, we obtain

$$u\left(A_t^* - \frac{A_{t+1}^*}{\alpha}\right) + \beta\delta V_{t+1}^* \geq u\left(A_t^* - \frac{X(A_t^*)}{\alpha}\right) + \beta\delta H(X(A_t^*)) \geq D\left(A_t^*\right), \tag{44}$$

where the second inequality follows from the fact that $\{X(A_t^*), H(X(A_t^*)\}$ is supportable at A_t^*. Because (44) holds for all $t \geq 0$, $\{A_t^*\}$ is an equilibrium path. ■

LEMMA **15.** *Suppose that for some $A^* \geq B$, $X(A) > A$ for all $A \geq A^*$. Then starting from any $A \geq A^*$, there is an equilibrium path with monotonic and unbounded accumulation, so that strong self-control is possible. Moreover, some such equilibrium path maximizes value among all equilibrium paths from A.*

Proof. We first claim that for any $A > A^*$ with $\lim_{A' \uparrow A} X(A') = A$, there is $\epsilon > 0$ with

$$X(A') = A \text{ for } A' \in (A - \epsilon, A). \tag{45}$$

Suppose on the contrary that there is $A > A^*$ and $\eta > 0$ such that $A' < X(A') < A$ for all $A' \in (A - \eta, A)$. Because $X(A) > A$, Lemma 14 and $\delta\alpha > 1$ together imply

$$H(A) > V^s(A) + \gamma \tag{46}$$

for some $\gamma > 0$.[44] Consider any equilibrium continuation $\{X(A'), V_1\}$ from $A' \in (A - \eta, A)$. Because $A'' < X(A'') < A$ for all A'' in that interval, $A_t' < A$ for the resulting equilibrium path. It follows from Lemma 9 that $V^s(A) > V_1$. Combining this inequality

[44]If $\delta\alpha > 1$ and $X(A) > A$, then the problem of Lemma 14 isn't solved by the stationary path from A: a small increase in assets followed by asset maintenance would achieve greater value.

with (46) and noting that $X(A') \to A$ as $A' \to A$,

$$u\left(A' - \frac{A}{\alpha}\right) + \beta\delta H(A) > u\left(A' - \frac{X(A')}{\alpha}\right) + \beta\delta V_1 \geq D(A')$$

for all $A' < A$ but close to A. So all such A' possess an equilibrium continuation of $\{A, H(A)\}$, which contradicts $X(A') < A'$, and establishes the claim.

We now complete the proof by claiming that any path $\{A_t\}$ from $A \geq A^*$ which solves the problem of Lemma 14 involves monotonic and unbounded accumulation. Suppose this assertion is false. Then at least one of the following must be true:

(i) there exists some date τ such that $A_\tau \geq A_{\tau+1} \leq A_{\tau+2}$, and/or

(ii) the sequence $\{A_t\}$ converges to some finite limit.

Let $\{c_t\}$ be the consumption sequence generated by $\{A_t\}$. In case (i), $c_\tau \geq c_{\tau+1}$. Recalling that $\delta\alpha > 1$, we therefore have

(47) $$u'(c_\tau) < \delta\alpha u'(c_{\tau+1}).$$

Moreover, because $X(A_\tau) > A_\tau$ and $A_\tau \geq A_{\tau+1}$, we have

(48) $$A_{\tau+1} < X(A_\tau).$$

In case (ii), there exists T such that, for $\tau > T$, (47) again holds because c_τ and $c_{\tau+1}$ are close. As far as (48) is concerned, there are two subcases to consider:

(a) There is $\tau > T$ with $A_{\tau+1} \leq A_\tau$. Here, (48) holds because $X(A_\tau) > A_\tau \geq A_{\tau+1}$.

(b) For $t > T$, A_t is strictly increasing with limit $\bar{A} < \infty$. If $\lim_{t\to\infty} X(A_t) > \bar{A}$, (48) plainly holds for some τ sufficiently large. Otherwise $\lim_{t\to\infty} X(A_t) = \bar{A}$. But in this case, we know from the first claim above that for some τ, $X(A_\tau) = \bar{A} > A_{\tau+1}$, so that (48) holds yet again for some τ sufficiently large.

In short, (47) and (48) always hold (for some τ). Now alter the path $\{A_t\}$ by increasing the period-$(\tau+1)$ asset level from $A_{\tau+1}$ to $A_{\tau+1} + \eta$, leaving asset levels unchanged for all other periods. Because $X(A)$ is non-decreasing, $A_{\tau+2} \leq X(A_{\tau+1} + \eta)$, and for small η we have $A_{\tau+1} + \eta < X(A_\tau)$ by (48); thus, the new path is feasible and also satisfies the constraints that define the value-maximizing path $\{A_t\}$. Taking the derivative of period-τ value with respect to η,

$$\frac{dV_\tau}{d\eta} = \delta^\tau \left[-u'(c_\tau)\frac{1}{\alpha} + \delta u'(c_{\tau+1}) \right] > 0,$$

where the inequality holds as a consequence of (47). This contradicts the definition of $\{A_t\}$ as a path that solves the problem in Lemma 14, and so establishes the lemma. ■

Proof of Proposition 2. Part (i) is obvious. "Only if" in part (ii) is also obvious, while "if" follows from Lemma 15. Likewise, the "only if" part of part (iii) is obvious, while the "if" part is a consequence of the fact that X is usc. Part (iv) once again is obvious. ■

Proof of Proposition 4, part (i). First suppose that there is $\epsilon > 0$ with $X(A) \geq A$ on $[B, B+\epsilon]$. By nonuniformity, $X(A') < A'$ for some A'. X is nondecreasing, so $X(S) = S$ for some $S > B$, with $X(A') < A'$ for some $A' \in (S, S+\epsilon')$, for every $\epsilon' > 0$.[45] By Lemma 11, S is sustainable. Define $\mu \equiv S/B$. By Lemma 8 (a), $\mu X(A'/\mu)$ is an equilibrium choice for every $A' \in [S, S+\mu\epsilon]$. But then $X(A') \geq \mu X(A'/\mu) \geq A'$ for all such A', a contradiction.

It follows immediately that $X(B) = B$, and for all $\epsilon > 0$, there exists $A_\epsilon \in (B, B+\epsilon)$ such that $X(A_\epsilon) < A_\epsilon$. But if the result is false, there is also $A'_\epsilon \in (B, A_\epsilon)$ with $X(A'_\epsilon) \geq A'_\epsilon$. Because $X(A)$ is nondecreasing, these observations imply the existence of $S_\epsilon \in (B, B+\epsilon)$ such that $X(S_\epsilon) = S_\epsilon$. By Lemma 11, S_ϵ is sustainable for all $\epsilon > 0$. But for ϵ sufficiently small,

$$D(S_\epsilon) \geq u\left(S_\epsilon - \frac{B}{\alpha}\right) + \beta\delta L(B) \geq u\left(S_\epsilon - \frac{B}{\alpha}\right) + \beta\delta V^s(B) > P^s(S_\epsilon)$$

where the first inequality follows from the definition of D, the second from Lemma 1, and the third from Lemma 13. This is a contradiction. ■

LEMMA **16** (Observation 3 in main text)**.** *Suppose that asset levels S_1 and S_2, with $S_1 < S_2$, are both sustainable, and that $X(A) > A$ for all $A \in (S_1, S_2)$. Then there exists $A^* \geq B$ such that $X(A) > A$ for all $A > A^*$.*

Proof. Let $\mu_i \equiv S_i/B$ for $i = 1, 2$; then $\mu_1 < \mu_2$. We claim that there is $A^* \geq B$ such that for all $A > A^*$, there are $\tilde{A} \in (S_1, S_2)$ and integers $(m, n) \geq 0$ with $A = \mu_1^n \mu_2^m \tilde{A}$.

[45] Take S to be the infimum of all A with $X(A) < A$.

We first show that there is A^* such that for all $A > A^*$, $A \in (\mu_1^k S_1, \mu_2^k S_2)$ for some k. Because $\mu_1 < \mu_2$, there is an integer ℓ with $\mu_1^{k+2} < \mu_2^{k+1}$ for all $k \geq \ell$. For all such k, $(\mu_1^k S_1, \mu_2^k S_2) = (\mu_1^k S_1, \mu_2^{k+1} B)$ overlaps with $(\mu_1^{k+1} S_1, \mu_2^{k+1} S_2) = (\mu_1^{k+2} B, \mu_2^{k+1} S_2)$. So $\cup_{k=\ell}^{\infty} (\mu_1^k S_1, \mu_2^k S_2) = \big(\mu_1^\ell S_1, \infty\big)$. Take A^* to be any number greater than $\mu_1^\ell S_1$.

Next we show that for each integer $k \geq 1$ and $A \in (\mu_1^k S_1, \mu_2^k S_2)$, there is $\tilde{A} \in (S_1, S_2)$ along with an integer $m \in \{0, \ldots, k\}$ such that $A = \mu_1^m \mu_2^{k-m} \tilde{A}$. Divide the interval $(\mu_1^k S_1, \mu_2^k S_2)$ (which is the same as the interval $(\mu_1^{k+1} B, \mu_2^{k+1} B)$) into a sequence of semi-open sub-intervals (preceded by an open interval) that coincide at their endpoints: $(\mu_1^{k+1} B, \mu_1^k \mu_2 B)$, $[\mu_1^k \mu_2 B, \mu_1^{k-1} \mu_2^2 B)$, $\ldots$, $[\mu_1 \mu_2^k B, \mu_2^{k+1} B)$. A must lie in one of these intervals; call it $[\mu_1^{m+1} \mu_2^{k-m} B, \mu_1^m \mu_2^{k-m+1} B)$, which we can rewrite as $[\mu_1^m \mu_2^{k-m} S_1, \mu_1^m \mu_2^{k-m} S_2)$. (The left edge is open if it is the first interval.) Thus, setting $\tilde{A} = A \mu_1^{-m} \mu_2^{m-k}$, we have $\tilde{A} \in (S_1, S_2)$ and $A = \mu_1^m \mu_2^{k-m} \tilde{A}$, as desired.

To complete the proof, pick any $A > A^*$ along with some $\tilde{A} \in (S_1, S_2)$, integer $k \geq 1$ and $m \in \{0, \ldots, k\}$ for which $A = \mu_1^m \mu_2^{k-m} \tilde{A}$. By repeated application of Lemma 8 (a), we see that $X(A) \geq \mu_1^m \mu_2^{k-m} X(\tilde{A})$; noting that $X(\tilde{A}) > \tilde{A}$, we have $X(A) > A$. ■

Let us refer to the assertion of Proposition 4, part (ii), as the Conclusion. Lemma 16 (together with Lemma 15) implies the Conclusion, provided that the supposition of Lemma 16 is satisfied. Via Lemma 16, several other situations also imply the Conclusion. Define $E(A) \equiv P^s(A) - D(A)$.

LEMMA **17.** $E(A) > 0$ *for some* $A > B$ *implies the Conclusion.*

Proof. Because u is continuous and D is increasing, there is an interval $[S_1, S_2]$ such that $E(A') > 0$ for all $A' \in [S_1, S_2]$ (e.g., take $S_2 = A$ and S_1 to be an asset level slightly below S_2). Clearly, S_1 and S_2 are both sustainable (indeed, every $A' \in [S_1, S_2]$ is).

For each $A' \in [S_1, S_2)$, define $z(A')$ as the largest value in $[S_1, S_2]$ satisfying

$$u\left(A' - \frac{z(A')}{\alpha}\right) + \beta\delta V^s(z(A')) \geq D(A'). \tag{49}$$

Because $E(A') > 0$, we have $z(A') > A'$. Moreover, because $E(z(A')) > 0$, we know that $z(A')$ is sustainable. So (49) and Lemma 3 imply the existence of an equilibrium starting from A' in which assets increase to $z(A')$ immediately and then remain at $z(A')$ forever. It follows that $X(A') \geq z(A') > A'$ for all $A' \in (S_1, S_2)$. Therefore the condition of Lemma 16 is satisfied: there are assets S_1 and S_2 with $S_1 < S_2$, both sustainable, with $X(A') > A'$ for all $A' \in (S_1, S_2)$. The Conclusion follows. ■

Say that a sustainable asset S is *isolated* if there is an interval around S with no other sustainable asset in that interval.

LEMMA **18.** *If S is sustainable and not isolated, then the Conclusion is true.*

Proof. Assume that S is sustainable and not isolated. By nonuniformity and Lemma 8, there is $A^* > S$ with $X(A^*) > A^*$. If $X(A') > A'$ for all $A' \geq A^*$, the Conclusion follows (Lemma 15). Otherwise, $X(A') \leq A'$ for some $A' > A^*$. Because X is nondecreasing, there is $S^* > A^*$ such that $X(S^*) = S^*$, and $X(A') > A'$ for all $A' \in [A^*, S^*)$.[46] By Lemma 11, S^* is sustainable.

Because S isn't isolated, for every $\epsilon > 0$ there is sustainable S' with $|S' - S| < \epsilon$. Let $\mu \equiv S/B$ and $\mu' \equiv S'/B$. By Lemma 8 (a), $S_1 \equiv \mu S^*$ and $S_2 \equiv \mu' S^*$ are sustainable. Remember that $X(A') > A'$ for all $A' \in [A^*, S^*)$. Using this information, it is easy to see that if S and S' are close enough, then $X(A) > A$ for all $A \in (S_1, S_2)$,[47] because all such A can then be written in the form $\mu' A'$ for some $A' \in (A^*, S^*)$. But now all the conditions of Lemma 16 are met, so that the Conclusion follows. ■

A special case of a sustainable asset level is what we will refer to as an *upper sustainable asset level* $\hat{S}$, one for which $X(\hat{S}) = \hat{S}$, while $X(A) > A$ over an interval of the form $[\hat{S} - \theta, \hat{S})$ for some $\theta > 0$. (Note that by Lemma 11, $\hat{S}$ is sustainable.)

LEMMA **19.** *Let $\hat{S}$ be upper sustainable. Then there is $\epsilon > 0$, such that for every $A \in [\hat{S}, \hat{S} + \epsilon]$, there is an equilibrium which involves first-period continuation asset $A_1 < \hat{S}$, and has value $V(A) < V^s(\hat{S})$.*

Proof. Using Lemma 13 and the fact that $\hat{S}$ is upper sustainable, there are $\zeta > 0$ and $\epsilon_1 > 0$ such that for every $A \in [\hat{S}, \hat{S} + \epsilon_1]$,

$$(50) \qquad u\left(A - \frac{\hat{S} - \zeta}{\alpha}\right) + \beta\delta V^s(\hat{S} - \zeta) \geq u\left(A - \frac{A_1}{\alpha}\right) + \beta\delta V^s(A_1)$$

[46]To see this, pick $S > A^*$ such that $X(S) = S$, and now take the infimum over all such values of S; call it S^*. Clearly, $S^* > A^*$ because $X(A^*) > A^*$ and X is nondecreasing.

[47]We presume that $S < S'$ without loss of generality.

whenever $A_1 \geq \hat{S}$, while at the same time,

(51) $$X(A'') > A'' \text{ for all } A'' \in [\hat{S} - \zeta, \hat{S}).$$

By part (i) of this proposition, there is $\tilde{A} > B$ such that every equilibrium from $A \in [B, \tilde{A})$ monotonically descends to B. By Lemma 8 (a) and the fact that $\hat{S}$ is sustainable, there must be a corresponding equilibrium which monotonically descends from A to $\hat{S}$ for every $A \in [\hat{S}, \hat{\mu}\tilde{A})$, where $\hat{\mu} = \hat{S}/B$. Define $\epsilon_2 \equiv \min\{\epsilon_1, \hat{\mu}\tilde{A} - \hat{S}\}$.

Using the first inequality in (34) of Lemma 9,

$$V^s(\hat{S}) \geq \sum_{t=0}^{\infty} \delta^t u\left(A_t - \frac{A_{t+1}}{\alpha}\right) + u'\left(\frac{\alpha - 1}{\alpha}\hat{S}\right)\left(\delta - \frac{1}{\alpha}\right)\zeta$$

for any path $\{A_t\}$ starting from $\hat{S}$ with the property that $A_t \leq \hat{S}$ for all $t \geq 0$, and $A_1 \leq \hat{S} - \zeta$. But then there exists $\epsilon_3 > 0$ such that

(52) $$V^s(\hat{S}) > \sum_{t=0}^{\infty} \delta^t u\left(A_t - \frac{A_{t+1}}{\alpha}\right)$$

for any path $\{A_t\}$ with $A_t \leq \hat{S}$ for all $t \geq 1$, $A_1 \leq \hat{S} - \zeta$, and $A_0 \leq \hat{S} + \epsilon_3$. Define $\epsilon \equiv \min\{\epsilon_2, \epsilon_3\}$.

Pick any $A \in [\hat{S}, \hat{S} + \epsilon]$, and consider any "descending equilibrium" as described just after (51), with payoff $P(A)$. Suppose that it has continuation (A_1, V_1). By Lemma 9, we know that $V_1 \leq V^s(A_1)$, so

(53) $$u\left(A - \frac{A_1}{\alpha}\right) + \beta\delta V^s(A_1) \geq P(A).$$

Combining (50) and (53), we must conclude that

(54) $$u\left(A - \frac{\hat{S} - \zeta}{\alpha}\right) + \beta\delta V^s(\hat{S} - \zeta) \geq P(A).$$

Now observe that (51), coupled with Lemma 14, implies that $H(\hat{S} - \zeta) \geq V^s(\hat{S} - \zeta)$. Using this information in (54), we must conclude that

(55) $$u\left(A - \frac{\hat{S} - \zeta}{\alpha}\right) + \beta\delta H(\hat{S} - \zeta) \geq P(A).$$

So the continuation $\{\hat{S}-\zeta, H(\hat{S}-\zeta)\}$ is an equilibrium from every $A \in [\hat{S}, \hat{S}+\epsilon]$. To complete the proof, note that any path $\{A_t\}$ associated with this equilibrium satisfies $A_t \leq \hat{S}$ for all $t \geq 1$,[48] $A_1 \leq \hat{S}-\zeta$, and $A_0 \leq \hat{S}+\epsilon \leq \hat{S}+\epsilon_3$. Therefore (52) applies. ■

Recall the definition of $d^*(A)$ as the largest maximizer of (20).

LEMMA **20.** *If* $d^*(A) = A$ *and* $d^*(A') \leq A'$ *over* $A' \in [A, A+\epsilon]$ *for some* $\epsilon > 0$, *then* A *is sustainable.*[49]

Proof. We first show that

$$L^+(A) \leq V^s(A). \tag{56}$$

By Lemma 5, L is increasing. So there is a sequence $\{A_n\}$ with $A_n \downarrow A$ and $L(A_n)$ (and $L^+(A_n)$) converging to $L^+(A)$. For each n, consider an equilibrium with the lowest value $V(A_n)$ among those that implement $Y(A_n)$.[50] Then

$$(1-\beta)u\left(A_n - \frac{Y(A_n)}{\alpha}\right) + \beta V(A_n) \geq D(A_n), \tag{57}$$

for all n. If strict inequality holds along a subsequence of n, then it's easy to see that $L(A_n) \leq V(A_n) = u(A_n - B/\alpha) + \delta L(B)$ along that subsequence.[51] Passing to the limit, $L^+(A) \leq u(A - B/\alpha) + \delta L(B) \leq V^s(A)$, where the second inequality comes from part (i) of the Proposition, already proved, which yields $L(B) = V^s(B)$, together with Lemma 9. So (56) holds in this case. In the other case, we may presume that

$$(1-\beta)u\left(A_n - \frac{Y(A_n)}{\alpha}\right) + \beta V(A_n) = D(A_n) \tag{58}$$

for all n. But in turn,

$$D(A_n) = u\left(A_n - \frac{d^*(A_n)}{\alpha}\right) + \beta\delta M(d^*(A_n), A_n). \tag{59}$$

---

[48]This follows from $X(\hat{S}) = \hat{S}$ and the fact that X is nondecreasing.

[49]In fact, a stronger property holds: if $d^*(A) \geq A$, then A is sustainable. That result follows directly from the existence of an everywhere-non-accumulating Markov-perfect equilibrium. Because we do not use the stronger property, nor do we focus on Markov equilibrium, we omit the proof.

[50]In line with Proposition 3, this value equals $L(A_n)$, but we do not use this fact anywhere in the proofs.

[51]We know that $Y(A_n)$ can be implemented by the continuation value $H(Y(A_n))$, and that it satisfies (35). If strict inequality holds in (35), reduce continuation assets, always using a continuation on the upper envelope of the value correspondence, and sliding down the vertical portion of H at any point of discontinuity. (Public randomization allows us to do this.) Note that payoffs and continuation values change continuously as we do this. Eventually we come to $Y(A_n) = B$ with continuation value $L(B)$.

Combining (58) and (59), we see that for every n,

$$(60)\quad (1-\beta)u\left(A_n - \frac{Y(A_n)}{\alpha}\right) + \beta V(A_n) = u\left(A_n - \frac{d^*(A_n)}{\alpha}\right) + \beta\delta M(d^*(A_n), A_n).$$

Now we pass to the limit in (60). By assumption, $d^*(A^n) \leq A^n$ for all n large, so $\lim_n d^*(A^n) < \alpha(1-\upsilon)A$.[52] By Lemma 7, d^* is right continuous at A, and so $d^*(A_n)$ converges to $d^*(A) = A$. By Lemma 6, $M(d^*(A_n), A_n) = L^+(d^*(A_n))$ for all n large enough, which converges to $L^+(d^*(A)) = L^+(A)$. Letting (Y, V) denote any limit point of $\{Y(A_n), V(A_n)\}$, we therefore have

$$(61)\qquad (1-\beta)u\left(A - \frac{Y}{\alpha}\right) + \beta V = u\left(\frac{\alpha-1}{\alpha}A\right) + \beta\delta L^+(A).$$

It follows that

$$\begin{aligned} \beta(1-\delta)L^+(A) &\leq \beta V - \beta\delta L^+(A) \\ &= u\left(\frac{\alpha-1}{\alpha}A\right) - (1-\beta)u\left(A - \frac{Y}{\alpha}\right) \\ (62)\qquad &\leq u\left(\frac{\alpha-1}{\alpha}A\right) - (1-\beta)u\left(\frac{\alpha-1}{\alpha}A\right) = \beta(1-\delta)V^s(A), \end{aligned}$$

where the first inequality uses $V(A_n) \geq L(A_n)$ for all n, so that $V \geq L^+(A)$, the equality follows from transposing terms in (61), and the second inequality uses $d^*(A_n) \geq Y(A_n)$ for all n, and $d^*(A_n) \to A$, so that $A \geq Y$. But (62) again implies (56).

With (56) in hand, we must conclude that

$$\begin{aligned} u\left(\frac{\alpha-1}{\alpha}A\right) + \beta\delta V^s(A) &\geq u\left(\frac{\alpha-1}{\alpha}A\right) + \beta\delta L^+(A) \\ &= u\left(\frac{\alpha-1}{\alpha}A\right) + \beta\delta M(A, A) \\ &= D(A) \end{aligned}$$

(where the last equality follows from $d^*(A) = A$), which means that A is sustainable. ■

In the rest of the proof, we make the assumption (by way of ultimate contradiction) *that the Conclusion is false*. Note that because many of the steps to follow are based on this presumption, they cannot all be regarded as relationships that truly hold in the model.

LEMMA **21.** *Suppose that the Conclusion is false. Then*

[52]That follows from $\alpha(1-\upsilon) > 1$, given $\alpha\delta > 1$ and $1-\upsilon > \gamma$, where γ is the Ramsey rate of saving.

(a) $d^*(\hat{S}) < \hat{S}$ *for any upper sustainable asset level* $\hat{S}$*, and*

(b) $d^*(A) \leq A$ *for all* $A \geq B$*, with strict inequality whenever* $X(A) \neq A$*.*

Proof. Part (a). Suppose not; then, since $X(\hat{S}) = \hat{S}$ (by the upper sustainability of $\hat{S}$), it follows from Lemma 4 that $d(\hat{S}) = \hat{S}$. We know that $M(\hat{S}, \hat{S}) = L^+(\hat{S})$ (see footnote 52 and recall Lemma 6), but by Lemma 19,

$$M(\hat{S}, \hat{S}) = L^+(\hat{S}) < V^s(\hat{S}).$$

Invoking (20) along with $d(\hat{S}) = \hat{S}$, we must therefore conclude that

$$D(\hat{S}) = u\left(\frac{\alpha - 1}{\alpha}\hat{S}\right) + \beta\delta M(\hat{S}, \hat{S}) < u\left(\frac{\alpha - 1}{\alpha}\hat{S}\right) + \beta\delta V^s(\hat{S}) = P^s(\hat{S}),$$

or $E(\hat{S}) = P^s(\hat{S}) - D(\hat{S}) > 0$. By Lemma 17, the Conclusion follows, a contradiction.

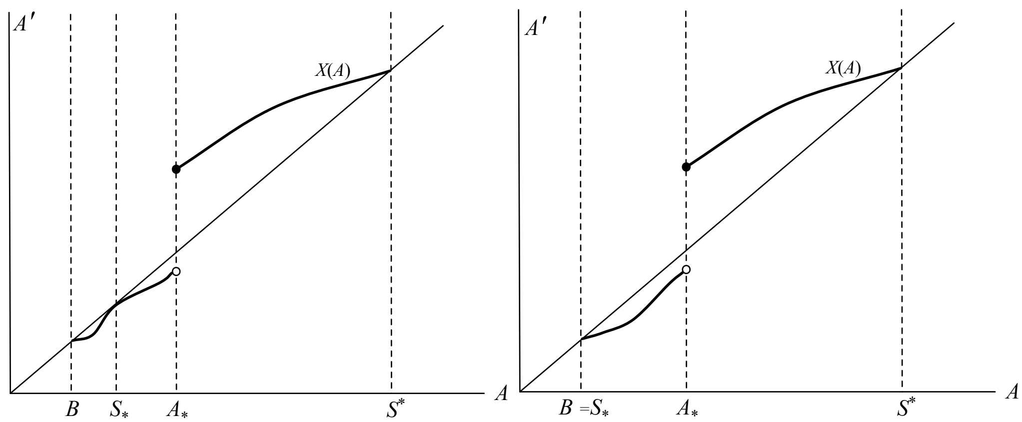
Part (b). If false, then $d^*(A) > A$ for some $A \geq B$, or $d^*(A) \geq A$ for some $A \geq B$ with $X(A) \neq A$. By Lemma 4, $X(A) \geq d^*(A)$, so in either case $X(A) > A$. Note that there is $A' > A$ such that $X(A') \leq A'$, otherwise Lemma 15 assures us that the Conclusion holds. Define $\hat{S}$ by the infimum value of such A'. Then it is immediate that $\hat{S}$ is upper sustainable, and that $X(A'') > A''$ for all $A'' \in [A, \hat{S})$.

Recall that $d^*(A) \geq A$, that d^* is nondecreasing and that $d(\hat{S}) < \hat{S}$ by the upper sustainability of $\hat{S}$ and part (a) of this lemma. So there is $S \in [A, \hat{S})$ with $d^*(S) = S$ *and* $d^*(S') \leq S'$ for all S' in an interval to the right of S.[53] By Lemma 20, S is sustainable.

Set $S = S_1$ and $\hat{S} = S_2$. Recall that $X(A'') > A''$ for all $A'' \in [A, \hat{S})$, so the inequality holds in particular on (S_1, S_2). Now all the conditions of Lemma 16 are satisfied. Together with Lemma 15, we see that the Conclusion must hold, a contradiction. ∎

Part (i) of the proposition, along with some of the foregoing lemmas, generates the following construction, *on the assumption that the Conclusion is false*. $X(A)$ starts out below A near B (there is a poverty trap by part (i)). By nonuniformity, $X(A) > A$ for some A; let A_* be the infimum value. $X(A) > A$ on an interval to the right of A_*; if not, sustainable stocks cannot all be isolated, and the Conclusion would follow

[53] To make this entirely clear, let $S \equiv \sup\{S' \in [A, \hat{S}) | d^*(S') > S'\}$. Because d^* is nondecreasing, $d^*(S) \geq S$. Moreover, $d^*(S) > S$ violates the definition of S (again, because d^* is nondecreasing).



(A) $S_* > B$ (B) $S_* = B$

FIGURE 7. THE TWO SUSTAINABLE ASSETS S_* AND S^*.

from Lemma 18.[54] Moreover, by Lemma 15, if the Conclusion is false, there is $S^* < \infty$, defined as the supremum of all asset levels S greater than A_* such that $X(A) > A$ for all $A \in (A_*, S)$. Note that S^* is upper sustainable. (Also note that $X(A_*) > A_*$, otherwise the Conclusion is implied by setting $S_1 = A_*$ and $S_2 = S^*$, and applying Lemma 16.)

Part (i) of the proposition also tells us that $d^*(B) = B$. Let S_* be the largest asset level in $[B, S^*]$ for which $d^*(S) = S$.

LEMMA **22.** *S_* is well-defined, with $B \leq S_* < S^*$, and $X(S_*) = S_*$.*

Proof. By Lemmas 18 and 21, there is a finite set of points in $[B, S^*]$, all strictly smaller than S^*, for which $d^*(S) = S$. (B is one such point.) So S_* is well-defined and $B \leq S_* < S^*$. That $X(S_*) = S_*$ follows from part (b) of Lemma 21 and $d^*(S_*) = S_*$. ■

Figure 7 summarizes the construction as well as the properties in Lemma 22. Panel A illustrates a case in which $S_* > B$, and Panel B, a case in which $S_* = B$. (Note: it is possible that $X(A) = A$ to the right of S_* and before S^*, though by Lemma 18, this can only happen at isolated points if the Conclusion is false.)

[54]By definition of A_*, there is $\{A'_n\}$ converging down to A_* with $X(A'_n) > A'_n$. If the assertion in the text is false, there is $\{A''_n\}$ also converging down to A_* along which $X(A''_n) \leq A'_n$. But then, using the fact that X is nondecreasing, there must be a third sequence along which equality holds, which proves that non-isolated sustainable assets must exist.

Define $Y^+(A)$ as the limit of $Y(A_n)$ as A_n converges down to A. Given Lemma 10, $Y^+(A)$ is well-defined and $Y^+(A) \geq Y(A)$.

LEMMA **23.** *If the Conclusion is false,* $Y^+(S_*) \geq S_*$.

Proof. If $S_* = B$ the result is trivially true, so assume that $S_* > B$. Suppose, on the contrary, that $Y^+(S_*) < S_*$. We first establish a stronger version of (56); namely, that

$$L^+(S_*) < V^s(S_*). \tag{63}$$

By part (b) of Lemma 21, $d^*(A) \leq A$ in a neighborhood to the right of S_* (indeed, strict inequality holds). With this in mind, carry out exactly the same argument as in the proof of Lemma 20, starting right after (56) and leading to (62), with S_* in place of A. We need two modifications to ensure that strict inequality in (56) holds. First, in case strict inequality holds in (57) along a subsequence, then $Y(A_n) = B$ and continuation values equal $L(B)$ along that subsequence, just as in the proof of Lemma 20, with the additional observation that (56) must indeed hold strictly, giving us (63). Otherwise, equality holds in (57), and (62) follows as before, with the additional implication that the second inequality in (62) — again, with S_* in place of A — must hold strictly, because $S_* > Y^+(S_*) \geq Y(S_*)$. We must therefore conclude that (63) holds, and therefore that

$$\begin{aligned} u\left(\frac{\alpha-1}{\alpha}S_*\right) + \beta\delta V^s(S_*) &> u\left(\frac{\alpha-1}{\alpha}S_*\right) + \beta\delta L^+(S_*) \\ &= D(S_*), \end{aligned}$$

where the equality follows from $d^*(S_*) = S_* < \alpha(1-\upsilon)S_*$, so that $L^+(S_*) = M(S_*, S_*)$ by Lemma 6. In other words, we have $E(S_*) > 0$. But then Lemma 17 assures us that the Conclusion must follow, which is a contradiction. ■

Let $\mu \equiv S^*/B$, and $\rho \equiv S_*/B$; then $\mu > \rho \geq 1$. Let $S_{**} \equiv \mu S_*$, and $S^{**} \equiv \mu S^*$. Note that $S_{**} = \mu S_* = \rho S^*$, so S_{**} is also a scaling of S^* by the factor ρ. (By Lemmas 11 and 22, S_* is sustainable, so Lemma 8 applies with both the scalings μ and ρ.)

Here is an outline of the remainder of the proof. Refer to Figure 8. By Lemma 8 (a), equilibria at assets to the right of S_* and to the left of S^* can be "scaled up" to assets beyond S_{**}, using the factor μ. Asset choices for such equilibria are partly indicated by the upper line to the right of S_{**} and the lower line to the left of S^{**}. But S_{**} is also a scaling of S^* (using ρ), so other equilibrium scalings are possible. In particular, Lemmas 8 and 19 tell us that equilibria with even lower values (and lower continuation assets)

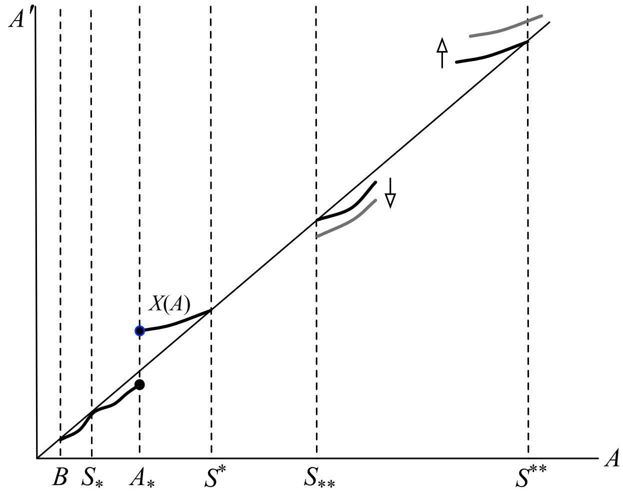


FIGURE 8. OUTLINE OF THE PROOF STARTING FROM LEMMA 24.

are achievable just above S_{**}; see the lower segment to the right of S_{**}. These values serve as punishments for deviations from even higher assets, and so support, in turn, larger asset choices near S^{**} relative to the earlier set of scaled equilibria; see the upper line around S^{**}. That creates a zone beyond S^{**} in which $X(A) > A$. If $X(A) > A$ for *all* $A > S^{**}$, Lemma 15 applies and the proof is complete. Otherwise, there is a first asset level beyond S^{**} at which $X(A) = A$ yet again. Now Lemma 17 applies, and contradicts the starting point of this entire construction: that the Conclusion is false.

Recall the definition of $L^+(x)$, and Lemma 6, which states that $M(x, A) = L^+(x)$ when $x < \alpha(1 - \upsilon)A$. This property will play a more active role now.

LEMMA **24.** *Suppose that the Conclusion is false.* (a) *For all* $x \geq B$*,*

$$L(\mu x) \leq \mu^{1-\sigma} L(x). \tag{64}$$

and in particular,

$$M(\mu x, \mu A) \leq \mu^{1-\sigma} M(x, A) \tag{65}$$

for all $A \geq B$ *and* $x \in [B, \alpha(1 - \upsilon)A]$.

(b) *For every* $A > S_*$ *with* $Y(\mu A) < S_{**}$ *and for all* $A' \in [S_*, A)$*,*

$$L^+(\mu A') < \mu^{1-\sigma} L^+(A'). \tag{66}$$

Proof. It is easy to see that Lemma 8 (a) implies (64). (65) follows for $x \in [B, \alpha(1-\upsilon)A)$ by taking right-hand limits of L, and for $x = \alpha(1-\upsilon)A$ by applying (64) directly. To prove part (b), pick $A > S_*$ with $Y(\mu A) < S_{**}$. Let $\tilde{A} \in (S_*, A]$. Because $Y^+(S_*) \geq S_*$ (by Lemma 23), any equilibrium from $\tilde{A}$ that implements $L(\tilde{A})$ has continuation $\{\tilde{A}_1, \tilde{V}_1\}$ with $\tilde{A}_1 \geq S_*$ (by Lemma 10). By Lemma 8 (a), $\{\mu\tilde{A}_1, \mu^{1-\sigma}\tilde{V}_1\}$ is an equilibrium continuation at $\tilde{A}'' \equiv \mu\tilde{A} > S_{**}$. So

$$u\left(\tilde{A}'' - \frac{\mu\tilde{A}_1}{\alpha}\right) + \beta\delta\mu^{1-\sigma}\tilde{V}_1 \geq D(\tilde{A}''), \tag{67}$$

and

$$\mu\tilde{A}_1 \geq \mu S_* = S_{**}. \tag{68}$$

Consider an equilibrium with the lowest continuation value — call this $\underline{V}$ — among those that implement $Y(\tilde{A}'')$ from $\tilde{A}''$. Then

$$u\left(\tilde{A}'' - \frac{Y(\tilde{A}'')}{\alpha}\right) + \beta\delta\underline{V} \geq D(\tilde{A}''). \tag{69}$$

If (69) does not bind, then we know that $Y(\tilde{A}'') = B$ and $\underline{V} = L(B)$ (see footnote 51). Recalling that $\tilde{A}'' = \mu\tilde{A}$, we must therefore have

$$\begin{aligned}
L(\mu\tilde{A}) &\leq u\left(\mu\tilde{A} - \frac{B}{\alpha}\right) + \delta L(B) \\
&\leq u\left(\mu\tilde{A} - \frac{\mu\tilde{A}_1}{\alpha}\right) + \delta\mu^{1-\sigma}\tilde{V}_1 - \frac{1-\beta}{\alpha\beta}u'\left(\mu\tilde{A} - \frac{B}{\alpha}\right)(\mu\tilde{A}_1 - B) \\
&\leq u\left(\mu\tilde{A} - \frac{\mu\tilde{A}_1}{\alpha}\right) + \delta\mu^{1-\sigma}\tilde{V}_1 - \frac{1-\beta}{\alpha\beta}u'\left(\mu A - \frac{B}{\alpha}\right)(S_{**} - B) \\
&= \mu^{1-\sigma}L(\tilde{A}) - \frac{1-\beta}{\alpha\beta}u'\left(\mu A - \frac{B}{\alpha}\right)(S_{**} - B), \qquad (70)
\end{aligned}$$

where the first inequality uses the definition of L, the second inequality uses Lemma 1, and the third inequality invokes (68) and $\tilde{A} \leq A$. On the other hand, if (69) does bind, then using (67) and noting that $\tilde{A}'' = \mu\tilde{A}$,

$$u\left(\mu\tilde{A} - \frac{\mu\tilde{A}_1}{\alpha}\right) + \beta\delta\mu^{1-\sigma}\tilde{V}_1 \geq u\left(\mu\tilde{A} - \frac{Y(\mu\tilde{A})}{\alpha}\right) + \beta\delta\underline{V}. \tag{71}$$

Let $\zeta \equiv S_{**} - Y(\mu A)$. Because Y is nondecreasing, we have $Y(\mu\tilde{A}) \leq S_{**} - \zeta \leq \mu\tilde{A}_1 - \zeta$. Using this information in (71) and observing that $\mu\tilde{A} \leq \mu A$, we must conclude that there exists $\eta_1 > 0$ with $\mu^{1-\sigma}\tilde{V}_1 \geq \underline{V} + \eta_1$, where η_1 might depend on A but can be chosen independently of $\tilde{A}$. Therefore, using (71) again, there is $\eta_2 > 0$ such that

$$u\left(\mu\tilde{A} - \frac{\mu\tilde{A}_1}{\alpha}\right) + \delta\mu^{1-\sigma}\tilde{V}_1 \geq u\left(\mu\tilde{A} - \frac{Y(\mu\tilde{A})}{\alpha}\right) + \delta\underline{V} + \eta_2,$$

or equivalently, $\mu^{1-\sigma}L\tilde{A}) \geq L(\mu\tilde{A}) + \eta_2$. Combining this inequality with (70) and defining $\eta \equiv \min\{\eta_2, [(1-\beta)/\alpha\beta]u'(\mu A - B/\alpha)(S_{**} - B)\}$, we have

$$\mu^{1-\sigma}L(\tilde{A}) \geq L(\mu\tilde{A}) + \eta \tag{72}$$

for all $\tilde{A} \in (S_*, A]$. Taking right-hand limits as $\tilde{A} \downarrow A' \in [S_*, A)$ in (72) then implies that $L^+(\mu A') < \mu^{1-\sigma}L^+(A')$ for all $A' \in [S_*, A)$. ■

LEMMA **25.** *Suppose that the Conclusion is false, and that for some $A \geq B$,*

$$L^+\left(d^*(\mu A)\right) < \mu^{1-\sigma}L^+\left(d^*(\mu A)/\mu\right). \tag{73}$$

Then

$$D(\mu A) < \mu^{1-\sigma}D(A). \tag{74}$$

Proof. By Lemma 21, $d^*(A') \leq A'$ for all $A' \geq B$, so by Lemma 6, $M(A', A') = L^+(A')$. Using this observation along with (73), we see that

$$\begin{aligned}
D(\mu A) &= u\left(\mu A - \frac{d^*(\mu A)}{\alpha}\right) + \beta\delta M(d^*(\mu A), \mu A) \\
&= \mu^{1-\sigma}u\left(A - \frac{d^*(\mu A)}{\mu\alpha}\right) + \beta\delta M(d^*(\mu A), \mu A) \\
&= \mu^{1-\sigma}u\left(A - \frac{d^*(\mu A)}{\mu\alpha}\right) + \beta\delta L^+\left(d^*(\mu A)\right) \\
&< \mu^{1-\sigma}\left[u\left(A - \frac{d^*(\mu A)}{\mu\alpha}\right) + \beta\delta L^+\left(\frac{d^*(\mu A)}{\mu}\right)\right] \\
&\leq \mu^{1-\sigma}\left[u\left(A - \frac{d^*(A)}{\alpha}\right) + \beta\delta L^+\left(d^*(A)\right)\right] \\
&= \mu^{1-\sigma}\left[u\left(A - \frac{d^*(A)}{\alpha}\right) + \beta\delta M\left(d^*(A), A\right)\right] = \mu^{1-\sigma}D(A),
\end{aligned}$$

where the second equality uses the constant-elasticity form of u, the strict inequality invokes (73), and the weak inequality follows from the definition of $d^*(A)$. ■

LEMMA **26.** *If the Conclusion is false,* $L^+(\mu A) < \mu^{1-\sigma} L^+(A)$ *for all* $A \in [S_*, S^*]$.

Proof. Because S^* is upper sustainable, Lemma 19 applies, so there is $\epsilon' > 0$ such that for every $A' \in (S^*, S^* + \epsilon']$, $Y(A') < S^*$. Because $S_{**} = \rho S^*$, Lemma 8 (a) implies that $Y(\rho A') < S_{**}$ for all such A'. In turn, this implies that for every $A'' \in (S_*, S_* + \epsilon]$, where $\epsilon \equiv \rho\epsilon'/\mu$, we have $Y(\mu A'') < S_{**}$. By part (b) of Lemma 24, $L^+(\mu A) < \mu^{1-\sigma} L^+(A)$ for all $A \in [S_*, S_* + \epsilon)$.

Suppose, by way of contradiction, that $L^+(\mu A) = \mu^{1-\sigma} L^+(A)$ for some $A \in [S_*, S^*]$. Let A^* be the infimum over such A. Then $A^* \geq S_* + \epsilon$ (by the conclusion of the last paragraph), and by the right-continuity of L^+,

$$L^+(\mu A^*) = \mu^{1-\sigma} L^+(A^*). \tag{75}$$

Define $A' \equiv \mu A^*$. There are now two cases to consider. First, if $d^*(A')/\mu > d^*(A^*)$,

$$\begin{aligned}
D(\mu A^*) = D(A') &= u\left(A' - \frac{d^*(A')}{\alpha}\right) + \beta\delta M(d^*(A'), A') \\
&= \mu^{1-\sigma} u\left(A^* - \frac{d^*(A')}{\mu\alpha}\right) + \beta\delta M(d^*(A'), A') \\
&\leq \mu^{1-\sigma}\left[u\left(A^* - \frac{d^*(A')}{\mu\alpha}\right) + \beta\delta M\left(\frac{d^*(A')}{\mu}, \frac{A'}{\mu}\right)\right] \\
&< \mu^{1-\sigma} D(A^*), \qquad (76)
\end{aligned}$$

where the weak inequality invokes (65), and the strict inequality the fact that $d^*(A^*)$ is the *largest* maximizer of $u(A^* - x/\alpha) + \beta\delta M(x, A^*)$, while $d^*(A')/\mu > d^*(A^*)$.

In the second case, $d^*(A')/\mu \leq d^*(A^*)$. Notice that (66) fails at $A = A^*$, so using part (b) of Lemma 24, $Y(\mu A) \geq S_{**}$ for all $A > A^*$. At the same time, $d^*(\mu A) \geq Y(\mu A)$ for all A (by Lemma 4). Combining these two observations, $d^*(\mu A) \geq S_{**}$ for all $A > A^*$.

By part (b) of Lemma 21, $d^*(\mu A) \leq \mu A$ for all A, so $\lim_{A \downarrow A^*} d^*(\mu A) \leq \mu A^* < \alpha(1-\upsilon)\mu A^*$. So Lemma 7 (b) applies, and d^* is right continuous at μA^*. Passing to the limit in the last inequality of the previous paragraph as $A \downarrow A^*$, it follows that $S_{**} \leq d^*(\mu A^*) = d^*(A')$, or $S_* \leq d^*(A')/\mu$. So in this second case,

$$S_* \leq d^*(A')/\mu \leq d^*(A^*) < A^*, \tag{77}$$

the last inequality following part (b) of Lemma 21, along with the fact that $A^* > S_*$, the latter being the largest value of $A \in [B, S^*]$ with $d^*(A) = A$.

In particular, (77) along with the definition of A^* allows us to verify condition (73) of Lemma 25 with A set equal to A^*. It follows that (74) holds at A^*. Recalling (76), we see then that in both cases

$$D(\mu A^*) < \mu^{1-\sigma} D(A^*). \tag{78}$$

Let $\{A_1, V_1\}$ be the equilibrium continuation that implements $L(A^*)$. By Lemma 8 (a), $\{\mu A_1, \mu^{1-\sigma} V_1\}$ is an equilibrium at μA^*, it has value equal to $\mu^{1-\sigma} L(A^*)$, and moreover, by the incentive constraint for $\{A_1, V_1\}$ coupled with (78),

$$u\left(\mu A^* - \frac{\mu A_1}{\alpha}\right) + \beta\delta\mu^{1-\sigma} V_1 \geq \mu^{1-\sigma} D(A^*) > D(\mu A^*).$$

This strict inequality, along with the fact that $\mu A_1 > B$, proves that one can lower equilibrium value at μA beyond the value created by scaling $\{A_1, V_1\}$, which shows that

$$L(\mu A^*) < \mu^{1-\sigma} L(A^*).$$

This contradicts the definition of A^*, and so completes the proof. ■

Proof of Proposition 4, part (ii). Assume the Conclusion is false. We claim that

$$E(S^{**}) = P^s(S^{**}) - D(S^{**}) > 0. \tag{79}$$

There are three possibilities to consider. First, $d^*(S^{**})/\mu \geq S_*$. We verify condition (73) of Lemma 25 with S^* in place of A. To do so, note that $d^*(S^{**})/\mu = d^*(\mu S^*)/\mu \geq S_*$, and also that $d^*(\mu S^*)/\mu \leq S^*$ by part (b) of Lemma 21. So we may apply Lemma 26 to $A = d^*(\mu S^*)/\mu$, and conclude that (74) is true for $A = S^*$. It follows that

$$D(S^{**}) < \mu^{1-\sigma} D(S^*). \tag{80}$$

Because $P^s(S^{**}) = \mu^{1-\sigma} P^s(S^*)$ and $P^s(S^*) \geq D(S^*)$, (80) immediately implies (79).

The second possibility is that $d^*(S^{**})/\mu < B$, so that $d^*(S^{**}) < \mu B = S^*$. Now apply part (b) of Lemma 8 by setting the path $\{\mu A_t^*\}$ in that lemma to the constant path with asset level $S^{**} = \mu S^*$ at every date.[55] It follows right away that $P^s(S^{**}) > D(S^{**})$, which establishes (79).

[55] This is our only use of part (b) of Lemma 8.

So the only remaining possibility is that

$$S_* > d^*(S^{**})/\mu \geq B. \tag{81}$$

Let d be a generic continuation asset choice that solves (20) at S^*. By Lemma 7 and the fact that $d^*(S_*) = S_*$, it must be the case that $d \geq S_*$. Because S^* is upper sustainable and so sustainable, and $d \geq S_* > d^*(S^{**})/\mu \geq B$, we see that if we define $A_1 \equiv d^*(S^{**})/\mu$, then

$$P^s(S^*) \geq D(S^*) > u\left(S^* - \frac{A_1}{\alpha}\right) + \beta\delta M(A_1, S^*). \tag{82}$$

Keeping in mind that $S^{**} = \mu S^*$ and $d^*(S^{**}) = \mu A_1$, we must conclude that

$$\begin{aligned} P^s(S^{**}) = \mu^{1-\sigma} P^s(S^*) &> \mu^{1-\sigma}\left[u\left(S^* - \frac{A_1}{\alpha}\right) + \beta\delta M(A_1, S^*)\right] \\ &= u\left(S^{**} - \frac{d^*(S^{**})}{\alpha}\right) + \beta\delta\mu^{1-\sigma} M(A_1, S^*) \\ &\geq u\left(S^{**} - \frac{d^*(S^{**})}{\alpha}\right) + \beta\delta M(d^*(S^{**}), S^{**}) \\ &= D(S^{**}), \end{aligned}$$

where the first inequality uses (82) and the second inequality uses (65). That gives us (79) again.

By Lemma 17, this immediately precipitates a contradiction, because (79) implies that the Conclusion follows, while we have been working with the presumption that the Conclusion is false. ∎

REFERENCES

Abreu, D. (1988), "Towards a Theory of Discounted Repeated Games," *Econometrica* **56**, 383–396.

Ainslie, G. (1975) "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control, *Psychological Bulletin* **82**, 463–496.

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

Akerlof, G. (1991). "Proscrastination and Obedience," *American Economic Review* **81**, 1–19.

Aliber, M. (2001), "Rotating Savings and Credit Associations and the Pursuit of Self-Discipline," *African Review of Money Finance and Banking*, 5172.

Aleem, I. 1993. "Imperfect Information, Screening, and the Costs of Informal Lending: A Study of a Rural Credit Market in Pakistan." In K. Hoff, A. Braverman and J. Stiglitz (eds.), *The Economics of Rural Organization: Theory, Practice and Policy*, chap. 7. London: Oxford University Press (for the World Bank).

Ambec, S. and N. Treich (2007), "Roscas as Financial Agreements to Cope with Self-Control Problems," *Journal of Development Economics* **82**, 120–137.

Anderson, S. and J-M. Baland (2002), "The Economics of Roscas and Intrahousehold Resource Allocation," *Quarterly Journal of Economics* **117**, 963–995.

Appadurai, A. (2004), "The Capacity to Aspire," in *Culture and Public Action* (V. Rao and M. Walton, eds), Stanford, CA: University Press.

Ariely, D. and K. Wertenbroch (2002), "Procrastination, Deadlines, and Performance: Self-Control by Pre-Commitment," *Psychological Science* **13**, 219–224.

Ashraf, N., Karlan, D. and W. Yin (2006), "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics* **121**, 635–672.

Bacchetta, P. and S. Gerlach (1997), "Consumption and Credit Constraints: International Evidence., *Journal of Monetary Economics* **40**, 207–238.

Banerjee, A. and A. Newman (1991), "Risk Bearing and the Theory of Income Distribution, *Review of Economic Studies/* **58**, 211–235.

Banerjee, A. and S. Mullainathan (2010), "The Shape of Temptation: Implications for the Economic Lives of the Poor," NBER Working Paper 15973, National Bureau of Economic Research.

Banerjee, A. and E. Duflo (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, PublicAffairs.

Basu, K. (2011), “Hyperbolic Discounting and the Sustainability of Rotational Savings Arrangements.” *American Economic Journal: Microeconomics* **3**, 143–71.

Bénabou, R. and J. Tirole (2004), “”Willpower and Personal Rules,” *Journal of Political Economy* **112**, 848–886.

Benartzi, S. and R. Thaler (2004), “Save More Tomorrow: Using Behavioral Economics to Increase Employee Savings,” *Journal of Political Economy* **112**, S164–S187.

Benhabib, J. and A. Bisin (2001), “Self-Control and Consumption-Saving Decisions: Cognitive Perspectives,” mimeo., Department of Economics, New York University.

Benoît, J-P. and V. Krishna (1985), “Finitely Repeated Games, *Econometrica* **53** 905–922.

Bernheim, B. D. and A. Dasgupta (1996), “Repeated Games with Asymptotically Finite Horizons, *Journal of Economic Theory* **67**, 129–152.

Bernheim, B.D., Ray, D, and S. Yeltekin (1999), “Self-Control, Saving, and the Low Asset Trap,” http://www.econ.nyu.edu/user/debraj/Papers/selfcontrol.pdf.

Bernheim, B. D. and A. Rangel (2009), “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” *Quarterly Journal of Economics* **124**, 51–104.

Beshears, J., Choi, J., Laibson, D. and B. Madrian (2011), “Self Control and Liquidity: How to Design a Commitment Contract,” Working Paper, Department of Economics, Harvard University.

Bryan, G., Karlan, D. and S. Nelson (2010), “Commitment Devices,” *Annual Review of Economics* **2**, 671–698.

DellaVigna, S. (2009), “Psychology and Economics: Evidence from the Field,” *Journal of Economic Literature* **47**, 315–372.

Fudenberg, D. and D. Levine (1983), “Subgame-Perfect Equilibria of Finite- and Infite-Horizon Games, *Journal of Economic Theory* **31**, 251–268.

Genicot, G. and D. Ray (2012), “Aspirations and Inequality.” Mimeo., Department of Economics, New York University.

Gine, X.,, Karlan, D. and J. Zinman (2010), “Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation,” *American Economic Journal: Applied Economics* **2**, 213–25.

Glick, R. and K. Lansing (2011), “Consumers and the Economy, Part I: Household Credit and Personal Saving,” *Federal Reserve Bank of San Francisco* Economic Letter, January 10; http://www.frbsf.org/publications/economics/letter/2011/el2011-01.html.

Goldstein, M. and C. Udry (1999), “Agricultural Innovation and Resource Management in Ghana,” mimeo., Yale Economics, http://www.econ.yale.edu/ cru2//pdf/finalrep.pdf.

Gugerty, M. (2007), “You Cant Save Alone: Commitment and Rotating Savings and Credit Associations in Kenya, *Economic Development and Cultural Change* **55**, 251–282.

Houser, D., Schunk, D., Winter, J. and E. Xiao (2010), “Temptation and Commitment in the Laboratory,” Working Paper No. 488, Institute for Empirical Research in Economics, University of Zurich.

Judd, K., Yeltekin, S. and J. Conklin (2003), “Computing Supergame Equilibria,” *Econometrica* **71**, 1239–1254.

Karlan, D., McConnell, M., Mullainathan, S. and J. Zinman (2010), “Getting to the Top of Mind: How Reminders Increase Saving,” Working Paper, Department of Economics, Yale University.

Karlan, D. and J. Zinman (2010), “Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts,” *Review of Financial Studies* **23**, 433–464.

Khera, R. (2011), “India’s Public Distribution System: Utilisation and Impact.” *Journal of Development Studies* **47**, 1038–1060.

Laibson, D. (1994), “Self-Control and Saving,” mimeo., Department of Economics, Harvard University.

Laibson, D. (1996), “Hyperbolic Discount Functions, Undersaving and Savings Policy,” NBER Working Paper w5635.

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

Ludvigson, S. (1999), “Consumption and Credit: A Model of Time-Varying Liquidity Constraints, *Review of Economics and Statistics* **81**, 434–447.

de Mel, S., McKenzie, D. and C. Woodruff (2008), “Returns to Capital in Microenterprises: Evidence from a Field Experiment,” *Quarterly Journal of Economics* **123**, 1329–1372.

Mookherjee, D. (1997), “Informational Rents and Property Rights in Land,” in *Property Rights, Incentives and Welfare,* J. Roemer (Ed.), Macmillan Press.

O’Donoghue, T. and M. Rabin (1999), “Doing It Now or Later,” *American Economic Review* **89**, 103–124.

Parker, J. (2000), “Spendthrift in America? On Two Decades of Decline in the U.S. Saving Rate, *NBER Macroeconomics Annual* **14**, 317–387.

Phelps, E. and R. Pollak (1968), “On Second-Best National Saving and Game Equilibrium Growth,” *Review of Economic Studies* **35**, 185–199.

Ray, D. (2006), “Aspirations, Poverty and Economic Change.” In A. Banerjee, R. Bènabou, and D. Mookherjee (eds), *What Have We Learnt About Poverty*, Oxford: Oxford University Press.

Shipton, P. (1992), “The Rope and the Box: Group Savings in the Gambia,” mimeo., Department of Anthropology, Boston University.

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

Thaler, R. (1990), “Anomalies: Saving, Fungibility, and Mental Accounts,” *Journal of Economic Perspectives* **4**, 193–205.

Thaler, R. and H. Shefrin (1981), “An Economic Theory of Self-Control,” *Journal of Political Economy* **89**, 392–406.

Udry, C. and S. Anagol (2006), “The Return to Capital in Ghana,” *American Economic Review* **96**, 388–393.

APPENDIX A. ALGORITHM

This section describes the iterative computational algorithm for obtaining an approximation to the equilibrium value correspondence $\mathcal{V}(A)$ through the sequence of correspondences $\{\mathcal{V}_k\}$ (See Section 3). Our initial correspondence is

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

in light of Observation 1.

The computational algorithm proceeds in four steps.[1] First, we consider a finite grid on the action and utility spaces. Second, given that continuation payoffs are governed by some correspondence $\mathcal{V}_k$, we determine the best-deviation payoffs at each asset level A (assuming the worst feasible punishments in the continuation set, which are well-defined given the discrete grid).

Third, we maximize and minimize value at each A subject to the no-deviation constraint and constraints on continuation utilities (that they be suitably drawn from $\mathcal{V}_k$). For this optimization step, we think of the individual as choosing the continuation level of assets rather than current consumption. This is convenient from a computational perspective.[2]

Finally, we use public randomization to construct $\mathcal{V}_{k+1}$ from the maximum and minimum values in Step 3, and test to see if convergence has occurred. The convergence criterion measures the largest difference (in the L^∞ norm) in utility bounds for each asset level between successive approximations. We end our iterations when this difference is "small," or more precisely, when

$$\max_{A\in\mathcal{A}}\{\max\{\ |L_k(A) - L_{k+1}(A)|, |H_k(A) - H_{k+1}(A)|\ \}\} < \epsilon$$

for some given precision parameter $\epsilon > 0$, where $\mathcal{A}$ is the discretized, finite action set from Step 1.

More formally, for a given set of parametric assumptions, our computational algorithm repeatedly applies the following four steps until convergence is achieved:

[1] This iterative numerical algorithm is a variation of the method of computing equilibria of supergames developed by Judd, Yeltekin and Conklin (2003).

[2] If consumption remains the choice variable, then we would need to discretize the consumption set. Additionally, the technology would have to be modified to ensure that for each current asset level and consumption choice, next period's assets are in the discretized asset set.

Step 1. *Initialization.*

1.1. Let $\mathcal{A}$ be a finite set of assets, chosen suitably fine and with a large upper bound.

1.2. Determine initial utility bounds $[L_0(A), H_0(A)]$ for each $A \in \mathcal{A}$.

Step 2. *Best Deviations.*

2.1. Let $\mathcal{A}(A_j) = \{A_i \in \mathcal{A} | A_j \geq c(A_i, A_j) \geq \nu A_j\}$ where $c(A_i, A_j) = A_j - A_i/\alpha$.

2.2. For each $A_i \in \mathcal{A}(A_j)$ compute

$$\tilde{D}(A_i, A_j) = \ u(c(A_i, A_j)) + \ \delta\beta L_k(A_i).$$

2.3. For each $A_j \in \mathcal{A}$ compute $D(A_j) = \max_{A_i \in \mathcal{A}(A_j)} \tilde{D}(A_i, A_j)$.

Step 3. *Highest and Lowest Values.*

3.1. Compute

$$H_{k+1}(A_j) = \max_{A_i \in \mathcal{A}(A_j)} \{u(c(A_i, A_j)) + \delta V_i\}$$

subject to the no-deviation constraint:

(a.1) $$u(c(A_i, A_j)) + \delta\beta V_i \geq D(A_j),$$

and the feasibility condition on continuation value:

(a.2) $$V_i \in \mathcal{V}_k(A_i).$$

3.2. Compute

$$L_{k+1}(A_j) = \min_{A_i \in \mathcal{A}(A_j)} \{u(c(A_i, A_j)) + \delta V_i\}$$

subject to exactly the same constraints (a.1) and (a.2).

Step 4. *Public Randomization and Convergence.*

4.1. Set $V_{k+1}(A) = [L_{k+1}(A), H_{k+1}(A)]$ (public randomization). Stop if convergence is reached; else return to Step 2.

Note that, in the maximization problem of Step 3.1, we must always set $V_i = H_k(A_i)$ as the continuation utility. After all, if any continuation value satisfies the no-deviation constraint (a.1), then so does the highest feasible continuation value, and that raises the overall value of the maximand as well. In contrast, in the minimization problem of

Step 3.2, we do not generally use $L_k(A_i)$ as the continuation utility, because the lowest feasible continuation value does not necessarily satisfy the no-deviation condition (a.1).[3]

For the results reported in Figure 1, we set $\sigma = 0.5$, so that

$$u(c) = \frac{1}{2}c^{1/2}.$$

Assets take on 8001 values between $[B, \bar{A}]$. We set $\bar{A} = 200$ and $B = 0.5$.[4] For the exercise depicted in Figure 1, we set the rate of return equal to 30%, the discount factor equal to 0.8, the hyperbolic parameter (β) equal to 0.4. Figure 1 Panel A plots the highest equilibrium asset choice, $X(A)$ and lowest equilibrium asset choice, $Y(A)$. Panel B plots the equilibrium value correspondence. For this particular exercise, a poverty trap exists below an asset level of 3.47. For initial asset levels above 3.47, however, there is indefinite accumulation.

Appendix B. Policy Regimes

In this section, we describe in more detail the extended model with taste shocks used in Section 6.3, as well as the policy regimes displayed in Figures 5 and 6. These regimes have a *lockbox* feature: assets are kept in an account with a rule specifying when and how much of the funds can be accessed. Each regime considers a different rule.

When $\alpha\delta > 1$, complete reliance on a lockbox always dominates internal rules provided that all consumption expenditures are perfectly foreseen; see discussion in main text. For these examples to have non-trivial solutions, we extend the original model to include an iid taste shock η (with probability distribution $p(\eta)$) that takes values in some finite set N and affects the flow utility in a multiplicative way. In every period, individuals make their saving/consumption decision after the realization of the current taste shock.

[3]However, Proposition 2 in the main text can be adapted to show that a carrot-and-stick structure obtains, so that often the highest continuation value (or some minor variant thereof) is also chosen in this problem.
[4]The analytical results allow for unbounded asset accumulation. An unbounded state space is not feasible computationally, but to ensure that the asset bound does not impact the policy and value functions reported in any significant way, we proceed in the following way. We choose an initial asset bound and note the asset level below this bound where the value and policy functions converge to the Ramsey solution ($\beta = 0$ case). We use the analytical Ramsey solution to approximate the value and policy functions beyond this intermediate asset value. We repeat this for a variety of intermediate asset values and initial asset bounds to check the robustness of the results for asset values below the intermediate asset level.

We first describe the baseline solution of this model without any lockboxes; it is a straightforward extension of solution with no taste shocks. Specifically, we can think of an *expected value correspondence* $\mathcal{V}^*(A; B)$ at the start of any date that defines the set of expected equilibrium values, the expectation taken over the taste shock which is about to be realized at that date, for every asset level. (For reasons that will become clear below, we explicitly carry the lower bound B, to be thought of as unchanging for all dates.) Because η is iid, $\mathcal{V}^*$ is the same at all dates. Thinking of these as continuation values from, say, date $t+1$, we can now define $\mathcal{V}^*(A, \eta; B)$ as the set of generated values at date t for any individual with asset level $A \geq B$, who has just experienced the taste shock η. The fixed-point logic of equilibrium generation then tells us that

$$\mathcal{V}^*(A; B) = \sum_{\eta \in N} p(\eta) V^*(A, \eta; B)$$

for every $A \geq B$, where we define the above convex combination of sets as the the collection of all elements that are themselves the same convex combinations of elements drawn from the individual sets.[5]

This value correspondence can be generated by a variation of the same iterated procedure described in Appendix A.

Now we consider regimes with lockboxes and thresholds. All the regimes we consider have the following lockbox properties: interest can always be withdrawn from the lockbox, which pays the same rate $\alpha - 1$ as a conventional savings account. No conventional savings is allowed until a threshold (A^T) is reached.[6] At that point, some or all of the lockbox principal is unlocked and made available. Let $\hat{B}$ denote the amount that still remains locked.

Recall that by convention, A includes non-financial labor income assets and an amount B is *always* "locked up" by the imperfect credit market. Therefore, we must constrain all our regimes by the property that $A^T \geq \hat{B} \geq B$.[7] In particular, we recover the standard problem by setting $A^T = \hat{B} = B$. Note that once past the threshold, the remainder of

[5]Under public randomization, each set is an interval and so all we need to do is convexity the best elements, and likewise the worst elements, and then draw the interval between these two numbers.

[6]The exercises we conduct are meant to be illustrative, and so we do not allow for contemporaneous savings while the lockbox is "active". These more realistic modifications can be easily studied, at least numerically.

[7]So, really, the *financial* assets in the lockbox are given by $A - B$, and all thresholds and locked amounts must be reinterpreted accordingly.

the problem facing the individual is *exactly* as in the standard case, without a lockbox feature, provided we replace the lower bound on assets by $\hat{B}$. So we can conceive of the overall problem as follows: at any date t, an individual is either "free" or "locked", depending on whether she has *ever* crossed the asset threshold A^T before date t. If she is free, then her (expected) value correspondence from that date onwards is governed by $\mathcal{V}^*(A, \hat{B})$. We can use this fact to anchor the construction of her value correspondence in the locked state. Denote this latter correspondence by $\hat{\mathcal{V}}$. It is to be noted that $\hat{\mathcal{V}}$ depends on the three parameters $(B, A^T, \hat{B})$, but we don't need to carry this dependence explicitly in the notation and so suppress it.

We can now determine best deviation payoffs (for every realization of the taste shock), as well as highest and lowest values, in the locked state. For every η and A in the locked state, consider the problem of finding

$$\hat{D}(A, \eta) \equiv \sup_{A' \in [A, \alpha(1-\upsilon)A]} \eta u \left(A - \frac{A'}{\alpha} \right) + \beta\delta L(A'), \tag{a.3}$$

subject to

$$L(A') = \begin{cases} \inf \mathcal{V}^*(A', \hat{B}) & \text{if } A' \geq A^T \\ \inf \hat{\mathcal{V}}(A') & \text{if } A' < A^T \end{cases} \tag{a.4}$$

Notice how the constraint in (a.3) requires $A' \geq A$: assets cannot be run down in the locked state. The second constraint describes where worst punishments following the deviation come from: if the choice of A' "frees" the individual, then it is drawn from the equilibrium value correspondence $\mathcal{V}^*(A', \hat{B})$ corresponding to the subsequent free state, and if the individual is still locked, it must come from the lowest value in $\hat{\mathcal{V}}(A')$. As a matter of fact, both infima in (a.4) can be shown to be attained, while in the discretized, finite computational problem under consideration, the "sup" in (a.3) can be replaced by "max".

With $\hat{D}$ in hand, we can turn to the problem of generating values at each A and η in the locked state. It is possible to generate any value V such that

$$V = \eta u \left(A - \frac{A'}{\alpha} \right) + \delta V'$$

for some A' with $A' \geq A$, and V' satisfying

$$V' \in \begin{cases} \mathcal{V}^*(A', \hat{B}) & \text{if } A' \geq A^T \\ \hat{\mathcal{V}}(A') & \text{if } A' < A^T \end{cases}$$

as long as the no-deviation constraint is also met:

$$\eta u\left(A-\frac{A'}{\alpha}\right)\beta\delta V' \geq \hat{D}(A,\eta).$$

Let $\hat{H}(A,\eta)$ and $\hat{L}(A,\eta)$ be the largest and smallest such values,[8] and recalling public randomization, define

$$\hat{\mathcal{V}}(A,\eta) \equiv [\hat{L}(A,\eta), \hat{H}(A,\eta)].$$

These are the "η-specific" value correspondences, and now we impose the fixed point consideration that

$$\hat{\mathcal{V}}(A) = \sum_{\eta\in N} p(\eta)\hat{\mathcal{V}}(A,\eta)$$

for every $A \in [B, A^T]$.

From a computational perspective, we discretize the space of assets and proceed exactly as in Appendix A to calculate $\hat{\mathcal{V}}$. That is, a two-stage procedure is employed, the first to determine the standard value correspondence $\mathcal{V}^*$ (for the lower bounds B and $\hat{B}$), followed by a similar process to obtain $\hat{\mathcal{V}}$. We omit the details here.

The text considers three regimes, all drawn from the class above. In Regime 1, plotted as a solid black line in Figures 5 and 6, the principal in the locked account is fully accessible after a specified $A^T > B$ is reached; so $\hat{B} = B$.

In Regime 2, shown as the dot-dash line in Figure 6, the threshold is eliminated. This corresponds to setting A^T equal to infinity in the above problem (the value of $\hat{B}$ is irrelevant). The individual can always withdraw current interest, but can never access the principal.

In Regime 3, corresponding to the dashed lines in Figure 6, contributions to the lock-up account stop once the threshold is reached, but the principal remains locked up forever. That is, $A^T = \hat{B} > B$. In this case, a switch to the standard problem occurs once the threshold is passed, but to a *different* standard problem, one characterized by the lower bound A^T on assets.

For the results displayed in Figures 5 and 6, the taste shock η takes two values, $\{0.8, 1.1\}$, with the associated probabilities $p(\eta = 0.8) = 0.3$ and $p(\eta = 1.1) = 0.7$. All other parameters are the same as in the earlier numerical results: the hyperbolic discount factor

[8] Once again, we disregard questions of attaining the maximum and minimum, which are trivial in the current finite context, but which can be affirmatively settled anyway.

(β) is 0.4, the geometric discount factor (δ) is 0.8, the constant elasticity parameter (σ) is 0.5, and B and $\bar{A}$ are set to 0.5 and 200 respectively. The standard problem with no lockbox features a poverty trap at low asset values. For $\eta = 0.8$, there is a poverty trap for $A < 4.42$ and for the high shock $\eta = 1.1$, a poverty trap exists when $A < 5.35$. For the first and third lock-up regimes, A^T is set to 5.5, slightly above the poverty threshold for the high taste shock state.