Focusing Effect and the Poverty Trap*

Andrea Canidio†

This Version: July 22, 2014.

Abstract

I build a dynamic consumption-savings model in which agents' choices are distorted by the *focusing effect*. I assume that agents overweight the utility of goods in which their options differ more. I show that the consumption-savings choice depends both on the marginal return on savings and on the *total return* on savings, implying that the incentive to save may increase with wealth. As a consequence, a salience-based poverty trap may exist when the marginal return on savings is sufficiently high and sufficiently flat. I also consider the case of perfect credit market and I show that a poverty trap may emerge when the salience of consumption is bounded above. I discuss policy implications. In particular, a punishment for decreasing savings below a threshold increases the salience of future consumption relative to present consumption and increases savings, even when this threshold is not binding.

JEL classification: D03, D31, O11, O15.

Keywords: Behavioral Poverty Trap, Salience, Focusing Effect, Inequality.

*I'm grateful to Adam Szeidl, Botond Kőszegi, Dilip Mookherjee, Andy Newman, Maitreesh Ghatak, Debraj Ray, Thomas Gall, Andrew Ellis and all participants to the Oslo ThRED workshop, SITE summer workshop 2014 (development economics), RES conference 2014 for their comments and suggestions.

†Department of Economics, Central European University, Nádor u. 9, 1051 Budapest, Hungary; email: canidioa@ceu.hu.

1 Introduction

Whether poverty and inequality are permanent conditions or should eventually disappear over time is a widely discussed issue in economics. At the center of the debate there are contrasting views on how the return on investment evolves with wealth. For example, in the neoclassical growth models developed in the 1950s (Solow-Swan model, Ramsey-Cass-Koopman model) inequality and poverty exist only in the short run and disappear over time. The basic assumption is that the marginal return on investment decreases with the size of the investment, so that a $1 investment produces a much higher percentage return than a $100,000 investment. This mechanism allows poor households to rapidly accumulate wealth, and to catch up with richer households over time. Irrespective of the initial conditions, all households eventually converge to the same level of wealth.

An alternative view, first proposed by Banerjee and Newman (1993) and Galor and Zeira (1993), is that the marginal return on investment may increase with wealth. For example, credit constraints may prevent poor agents from borrowing and therefore may exclude them from projects that require a high initial investment. Under this assumption, identical households starting with different wealth levels may converge to different steady states. Inequality and poverty are determined by the initial conditions and can be persistent. Policy interventions can have dramatic effects. For example, a one-off wealth transfer may permanently eliminate poverty.

However, convincing empirical evidence shows that poor households have very high marginal-returns projects available, but often decide not to invest. Poor households routinely forgo investments such as buying a mosquito net, vaccinating their children (see Banerjee and Duflo, 2011), using fertilizers (see Duflo, Kremer, and Robinson, 2011), switching to more valuable crops (see Udry and Anagol, 2006), or keeping spare change for their small businesses (see Beaman, Magruder, and Robinson, 2013). Most of these investments yield returns of 50-100% and are divisible (e.g. fertilizer). This evidence raises the suspicion that poor households remain poor not because they lack profitable investment opportunities, but because behavioral biases prevent them from exploiting these opportunities. In particular, note that these high marginal-return investments produce extremely low *total* returns because poor households can invest very little.

It is well known that the *focusing effect* can cause people to overlook small gains. The focusing effect (or focusing illusion) occurs whenever an agent places too much importance on certain aspects of her choice set or on certain pieces of information (i.e. certain elements are more *salient* than are others). For example, Schkade and Kahneman (1998) show that, when asked about comparing life in California and in the Midwest, most people report California as the best place to live and cite the weather - i.e. the dimension in which the two choices differ the most - as the main reason. Despite this, actual measures of life satisfaction in the two regions are similar, implying that the Midwest is better than California in some dimensions other than the weather. However, the difference between the two regions in each of these other dimensions is too small to be salient, causing people to ignore them when evaluating the two options.[1]

In this paper, I introduce the focusing effect as modeled by Kőszegi and Szeidl (2013) into a dynamic consumption-savings model. I show that the focusing effect distorts the consumption-savings decision, and that the severity of this distortion depends on the agent's initial wealth. I show that a *salience*-induced poverty trap is possible. In the model, I assume that the marginal return on investment is arbitrarily large for a very small investment and arbitrarily small for a very large investment, as in the neoclassical growth models. Nonetheless, a poverty trap may emerge because the salience of future consumption depends on the *total* return on savings. Hence, a $1 investment with a 100% return is less salient than a $100,000 investment with a 1% return. The total return on investment increases with wealth, increasing the salience of future consumption and the incentive to save as wealth increases. At the same time, the marginal return on investment decreases with the size of the investment, decreasing the incentive to save as wealth increases. Overall, the incentive to save may increase or decrease with wealth.

The existence of a poverty trap is determined by the level of the marginal return on savings, and by how rapidly it decreases with wealth. If the marginal return on savings is very high and decreases slowly with wealth, then the *total* return on investment increases rapidly with wealth and a poverty trap is possible. In this case, the salience of future consumption relative to current consumption increases with wealth, and salience is the main determinant of the incentive to save at different wealth levels. In other words, a poverty trap exists because high-return projects are available to poor people and to those who are slightly less poor. For example, if fertilizer delivers the same percentage return at different investment levels, then a poverty trap is possible. A poor agent and a relatively richer agent both invest in fertilizer, but the total return on investment is lower for the poor agent. It follows that the salience of savings is lower for the poor agent than for the richer one. Because the marginal return on investment is constant for both agents, differences in the incentive to save are driven solely by differences in the salience of savings. These different initial incentives to save may lead, in the long run, to different steady-state wealth levels.

The fact that people's preferences may depend on the choice set available is well documented. For example, when choosing between two vectors of goods x and y, a person may pick x when the choice set is $\{x, y\}$ and may pick y when the choice set expands to $\{x, y, z\}$. This violation of rationality is relevant for understanding economic development because economic development leads to an expansion in the available choices. The psychological literature identified the focusing effect as one of the reasons why preferences may change with the choice set, arguing that the introduction of a new element in the choice set may affect the salience the other elements of the choice set. Kőszegi and Szeidl (2013) formalize this concept by assuming that agents evaluate each

[1] Similarly, Kahneman, Krueger, Schkade, Schwarz, and Stone (2006) show that people place too much weight on differences in monetary compensation when asked to compare job offers. The interpretation is that the non-monetary component of each job offer does not vary much across job offers and is therefore not salient.

bundle using a *focus-weighted utility*

$$\tilde{U}(x_1, x_2, .., x_n) = \sum_{s=1}^{n} g_s u_s(x_s)$$

where $x_1, x_2, .., x_n$ are different goods in bundle x. The weight g_s is the *focus weights* attached to good s, which represents the salience of good s and is defined as:

$$g_s = g\left(\max_{x \in C} u_s\left(x_s\right) - \min_{x \in C} u_s\left(x_s\right)\right)$$

where C is the choice set and $g()$ is the *focus function,* assumed strictly increasing. In this formalization, agents overweight the utility of goods in which their options differ more, when these differences are given by the maximum and minimum utility that is possible to achieve by consuming a given good. In turns, the salience of a good affects the sensitivity of the agent to the utility provided by that good.

Kőszegi and Szeidl (2013) argue that the focusing effect leads to a *bias toward concentration*, i.e. the choice maker overvalues concentrated costs (or benefits) relative to disperse benefits (or costs). This bias explains, for example, why most taxpayers prefer one lump-sum tax refund over monthly withholdings, why most retirees take too much of their retirement wealth in a lump sum rather than as an annuity, why financing for consumer products is so popular. For the purpose of this paper instead, the focusing effect is relevant because as wealth increases new consumption bundles become available, which are better than the ones previously affordable and therefore change the salience of all choices available. Whether present consumption bundles or future consumption bundles becomes relatively more salient depends the utility of the newly available future consumption bundles relative to the utility of the newly available present consumption bundles.

I build a model in which agents live for two periods. In the first period, they consume and save. In the second period, they consume and leave bequests. Agents' utility depends on their consumption and on the bequests left to the following generation. The utility from present and future consumption and the utility from bequests are discounted by focus weights, which depend on the wealth level at the beginning of life and on the return on savings. For example, if the return on savings is high, the future will be salient because the maximum level of future achievable consumption is high. Hence the salience of present consumption, future consumption and bequests are increasing with wealth, but the rates at which the different focus weights grow depend on the return on savings.

The focusing effect introduces a wealth-dependent discount factor in the form of a *focus wedge*, which is defined as the salience of future consumption relative to present consumption. I derive sufficient conditions for the focus wedge to increase with wealth, involving the shape of the focus function, the utility function, and the return on investment. I show that multiple steady states are possible when the marginal return on investment is flat but high. Intuitively, the focus weight increases with wealth if technology is productive enough, because as wealth increases future consumption possibilities expand faster than present consumption possibilities. At the same time, the incentive to save depends also on the marginal return on savings, which decreases with wealth. Overall, if the marginal return is relatively flat and high, multiple steady states may exist. Hence, the observation that poor people have high marginal return project available is consistent with the existence of a poverty trap, provided that the the same high-marginal return projects are available at different wealth levels.

The model has a unique per-period equilibrium, so that different steady states are reached depending on the initial condition of the economy. However, the savings function needs not to be monotonic: sometimes wealthier agents save less than poorer agents because the focus wedge may be locally decreasing with wealth. It follows that next to poverty traps, the model may generate *middle-income traps*, in which households starting with low and high wealth levels converge to the same high steady-state wealth level, while households starting with intermediate wealth levels converge to a lower steady-state wealth level.

I also consider the case of perfect credit market by assuming that agents can freely borrow and lend at a given interest rate. I show that a poverty trap is possible also in this case. The driving force is the shape of the utility function. If the utility function is bounded above, the salience of consumption (present or future) is bounded above. Hence, as wealth increases the difference between the salience of present consumption and the salience of future consumption decreases sufficiently fast. The distortion in the consumption-saving decision introduced by the focusing effect is less relevant for rich agents than for poor agents, implying that rich agents save more than poor agents do.

The first obvious policy implication is that a one-off wealth transfer may have long-term effects because it may push the economy to a higher steady state. A second, more interesting, aspect of the model is that any policy that alters the agent's choice set even temporarily may have a permanent effect, irrespective of whether the agent chooses any of the new options introduced by the policy or whether the policy constraints the agent's choice. For example, a nonlinear subsidy to savings may increase the savings rate even if the agent does not benefit from the subsidy. A one-off tax on consumption may make savings relatively more salient, even if the agent does not bear the cost of the tax.

Finally, I allow agents to manipulate their choice set and future preferences using a commitment-saving device, which imposes a punishment if the stock of savings drops below a given threshold. This type of commitment devices are widespread and commonly used in the developing world. For example, in ROSCAS the penalties for not saving enough are harsh social sanctions and social exclusion. Similarly, poor households save by borrowing at very high interest rates from money lender and MFIs, who deliver a punishment if there is no repayment.[2] The typical explanation for the value of commitment relies on hyperbolic discounting (see Laibson, Repetto, and Tobacman, 2003). Here instead, commitment decreases the salience of present consumption, increases the value of future consumption and allows for asset accumulation. Note that a model in which choices are distorted by the focusing effect delivers a specific empirical prediction: commitment-savings devices should have an impact on savings *even when they are non-binding*. This implies that, for example, agents who adopt commitment-savings devices may increase their savings above the minimum amount imposed by the commitment device.

1.1 Relevant Literature

In the economic literature, the focusing effect is called *salience*, *focusing*, and *attention* (or *inattention*.) Different strands of literature use different words, reflecting different ways of formalizing this concept.

One branch of literature proposes models in which the focusing effect is embedded directly into the agent's preferences. For my purposes, the most relevant works using this framework are Kőszegi and Szeidl (2013) (who call this distortion *focusing*), and Bordalo, Gennaioli, and Shleifer (2013) (who call this distortion *salience*). Kőszegi and Szeidl (2013) assume that the variation in utility levels achievable within the choice set affects the agent's choice, and pushes the agent to overvalue the goods (or goods attributes) that vary the most within the choice set. Bordalo, Gennaioli, and Shleifer (2013) assume that agents overvalue the goods that differ the most with respect to a reference point. Kőszegi and Szeidl (2013) is well suited to analyze how wealth accumulation affects the salience of present and future consumption, because wealth accumulation leads to an expansion of the choice set. Performing the same analysis using the Bordalo, Gennaioli, and Shleifer (2013) approach requires first to establish how wealth accumulation affects the agent's reference point.

The second strand of literature models the focusing effect as a rational choice given some information-absorption constraint. In other words, agents know that they can use only a given "amount of information", so they choose strategically what variable to consider when making a decision (hence the name *rational inattention*, see Sims, 2003 or a recent paper by Gabaix, 2012.) Hence, agents do not put attention on goods (or goods attributes), but on pieces of information.

The last strand of literature models attention as a limited resource that can be allocated to different problems. For example, Banerjee and Mullainathan (2008) show that limited attention can generate a poverty trap. In their model, attention is a scarce resource that can be employed either in production (where it reduces mistakes) or at home (where it solves problems). Crucially, attention and home consumption are substitutes at home, while attention and human capital are complements at work. High-human capital agents devote attention at work and generate high

[2] Ananth, Karlan, and Mullainathan (2007) document that in several parts of the world microentrepreneurs borrow at daily rates of around 10%, for several years in a row.

income, while low-human capital agents devote attention at home generating low income. If income level and human capital of offspring are correlated (e.g. because of credit market imperfections), then this mechanism delivers an attention-based poverty trap. In the model presented here attention is not an input in the production functions, but it enters directly into the utility function. However, I show in the appendix that my model can be interpreted as the reduced form of a costly attention model. The key assumption is that the cost of attention depends on the size of the choice set, and that as the choice set changes the agent optimally puts more attention to the dimension that changed the most.

Other authors have shown that behavioral biases are worsened by poverty, therefore generating behavioral poverty traps. Bernheim, Ray, and Yeltekin (2013) analyze a growth model with time-inconsistent agents. They consider all possible equilibrium consumption-savings paths, and show that the one with highest wealth accumulation increases with wealth. Agents can commit to a high level of savings by employing a *personal rule*: if any past self deviated from the equilibrium, then the present self consumes as much as possible (i.e. as much as allowed by an equilibrium). The punishment that follows a deviation therefore increases with wealth, meaning that equilibria with higher investment rates become sustainable as wealth increases.

Few authors argued that behavioral biases may lead to a convex savings function and poverty traps.[3] Banerjee and Mullainathan (2010) analyze a consumption-savings problem in which different goods have a different discount factor. Some goods are *temptation goods*: the present discounted value of their future consumption is low (or even zero); nonetheless they will be consumed in the future. The presence of temptation goods creates a discount between present and future. Banerjee and Mullainathan (2010) show that if the share of temptation goods consumed as a fraction of wealth decreases with wealth, then poorer people discount the future more than rich people do. Thus a temptation-based poverty trap may exist. Moav and Neeman (2012) develop a theory in which conspicuous consumption and the concern for social status may lead to a poverty trap. In their model, people infer the social status of other people by observing their human capital investment and expenditure in conspicuous consumption. Hence, concerns for status acts as a regressive tax, affecting poor and low-human capital agents more than rich and high-human capital agents.[4]

My paper illustrates a very different type of behavioral poverty trap, and it complements the work of Banerjee and Mullainathan (2008), Banerjee and Mullainathan (2010), Moav and Neeman (2012) and Bernheim, Ray, and Yeltekin (2013). In particular, here a poverty trap exists when preferences are distorted by the focusing effect and when the investment opportunity frontier has a specific shape.

The remainder of the paper is structured as follows. In section 2 I derive a simple two-periods consumption-savings model. In section 3 I expand the simple two-periods model to an infinite

[3] See Moav (2002) for a treatment of how convex savings function can generate poverty traps.
[4] See also Moav and Neeman (2010).

horizon, and I shows that a poverty trap is possible. In section 4 I describe the dynamics of the model. In section 5 I consider the case of perfect credit markets. In section 6 I analyze the effect of commitment savings devices. Section 7 concludes.

2 Preliminary: Focusing Effect in a Two-Periods Consumption-Savings Model

In this section I build a simple two-periods consumption-savings model with focusing effect, and I provide some intuition for the main results of the paper. All assumptions and formal derivations are postponed to the next sections.

Consider the following consumption-savings problem with focusing effect:

$$\max_{c_1, c_2} \{h(b)u(c_1) + h(f(b))u(c_2)\}$$

$$s.t \;\; f(b - c_1) = c_2$$

where b is initial wealth, c_1 and c_2 are consumption levels in the two periods of life. The production function $f(.)$ determines the return on savings and represents the outer envelope of the return on all investment opportunities that are available to the agent.

The focusing effect enters the above problem via the *focus weights* $h(b)$ and $h(f(b))$, defined as

$$h(x) \equiv g\left(u(x) - u(0)\right)$$

where $g(.)$ is the *focus function* assumed strictly increasing. In other words, the salience of consumption in a given period depends on the maximum and on the minimum utility achievable in that period. The minimum utility achievable is always $u(0)$. The maximum utility achievable depends on the maximum consumption achievable, which in period 1 is b (in case the agent does not save anything), and in period 2 is $f(b)$ (in case the agent saves everything). In turns, the salience of present and future consumption determines the sensitivity of the agent to the utility enjoyed in the present and in the future, and distorts the consumption-savings choice. Note that the salience of consumption in both periods increases with wealth. However, as wealth increases, the salience of period-1 consumption and the salience of period-2 consumption will, in general, grow at different rates.

The first order condition of the problem is

$$u'(c_1) = f'(b - c_1)\left[\frac{h\left(f(b)\right)}{h(b)}\right]u'(c_2)$$

The focusing effect distorts the savings decision via the *focus wedge* $\frac{h(f(b))}{h(b)}$. Through the focus wedge, for given b, the *total return on investment* affects the salience of future consumption and the consumption-savings decision. More interestingly, the shape of the focus wedge determines how the incentive to save changes as initial wealth increases. I provide all formal derivations in the next sections (see lemmas 3, 4 and 5), but it is already possible to see that the shape of the focus wedge depends on:

- The production function. The fastest $f(x)$ grows (highest $f'(x)$) the more likely is $\frac{h(f(b))}{h(b)}$ to be increasing with b. Note that when the marginal return on savings is particularly high for small investment levels (as in the neoclassical growth models), the focus wedge is likely to be increasing in b for low levels of initial wealth.
- The shape of $u(x)$ and the shape of $g(x)$. If either $u(x)$ or $g(x)$ are bounded above, then the focus wedge $h(.)$ is also bounded above. This implies that, as wealth increases, the difference in salience between future and present becomes smaller.

Overall, changes in initial wealth affect the incentive to save in two ways. On the one hand, if the focus wedge is increasing in b, future consumption may become relatively more salient than present consumption. On the other hand, the marginal return on investment decreases with wealth, decreasing the incentive to save. I show formally later that if the marginal return on investment is approximately constant, then the first effect dominates. In this case salience is the first-order determinant of the incentive to save, which can be increasing with wealth.

In the next section, I modify the above model by assuming that each generation leaves a bequests to the following generation, and that the size of the bequest left is correlated to amount saved between period 1 and period 2. In that model, because of salience, bequest left as a share of bequests received may be locally increasing, implying that a salience-based poverty trap may exist.

3 Focusing Effect in an Infinite-Horizon Consumption-Savings Model

Consider the infinite-horizon version of the problem described above. An agent is born at time t and lives for two periods. In the first period she consumes and decides how much to save. In the second period she consumes and decides how much to leave as bequests to her offspring. The problem faced by this agent is:

$$\max_{c_{1,t},c_{2,t},b_{t+1}} \{h_1(b_t)u_1(c_{1,t}) + h_2(c^\star(f(b_t)))u_2(c_{2,t}) + h_b(f(b_t) - c^\star(f(b_t)))\nu(b_{t+1})\}$$

$$s.t \;\; f(b_t - c_{1,t}) = c_{2,t} + b_{t+1}$$

Bequests received are denoted by b_t, while bequests left are denoted by b_{t+1}. Consumption by the agent born in period t during the period of life i is denoted by $c_{i,t}$, while $c^\star(x)$ is the amount that will be consumed in the second period of life if the agent saves x. The functions $u_1(.)$, $u_2(.)$ and $\nu(.)$

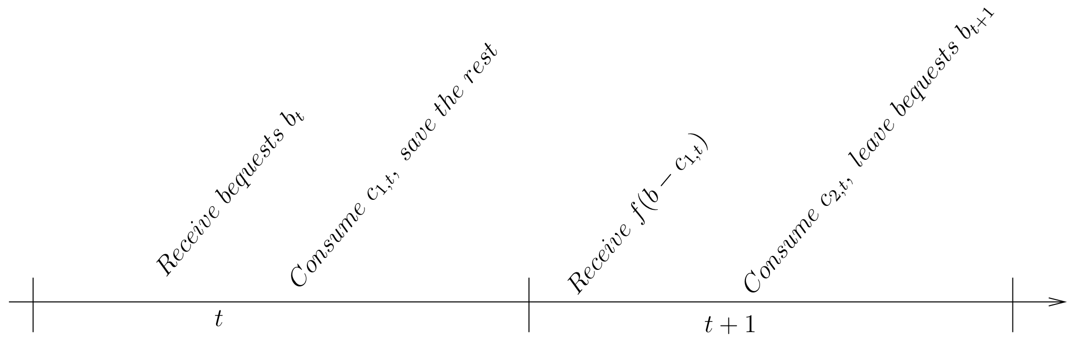


Fig. 1: Timeline

are assumed strictly concave, strictly increasing, continuous, differentiable with $u_1(0)$, $u_2(0)$, $\nu(0)$ finite. The production function $f(.)$ is assumed continuous, differentiable, strictly increasing, strictly concaves, unbounded above, with $f(0) = 0$, $\lim f'(x) = \infty$ and $\lim f'(x) = 0$.

The focus weights are

$$h_1(x) \equiv g\left(u_1(x) - u_1(0)\right)$$

$$h_2(x) \equiv g\left(u_2(x) - u_2(0)\right)$$

$$h_v(x) \equiv g\left(u_v(x) - \nu(0)\right)$$

Also here, the salience of period-1 consumption, period-2 consumption, and bequests depend on the maximum utility achievable from the consumption of each good. However, here the agent cannot decide in period 1 to allocate all his wealth to bequests or to period-2 consumption. In period 1, the agent can only decided how much to consume and to save, and the agent's future self will then decide on how to split savings between bequests and period-2 consumption. Therefore, the salience of future consumption and future bequests depend on the consumption decision of period-2 self in case period-1 self saves all the available wealth, and are given by $h_2(c^\star(f(b_t)))$ and $h_b(f(b_t) - c^\star(f(b_t)))$.

Assumption 1. *$u_2(x) = \nu(x)$ for all x.*

By assuming that $u_2(x) = \nu(x)$ the problem can be rewritten as

$$\max_{c_t, b_{t+1}} \left\{ h_1(b_t) u_1(c_t) + h_2\left(\frac{f(b_t)}{2}\right) 2 \cdot u_2(b_{t+1}) \right\}$$

$$s.t \;\; f(b_t - c_t) = 2b_{t+1}$$

where I used the fact that, in period 2, savings will be equally split between consumption and bequests. Hence, under assumption 1, b_{t+1} is a simple linear function of the return on savings. In addition, I can simplify the notation and call c_t the consumption in the first period of life, and call b_{t+1} both bequests left and consumption in the second period of life.

Assumption 2. *$g(x)$ is strictly increasing, strictly concave, continuous and differentiable, with $g(0) > 0$.*

The concavity of $g(x)$ can be related to the Weber-Fechner law of human perception: the intensity of a sensation is proportional to the logarithm of the intensity of the stimulus causing it. In this case, the focus weight measures the intensity of the utility of consumption. Weber-Fechner law implies that the focus weight should be proportional to logarithm of utility. Here I make a more general assumption and I simply impose that $g(x)$ is concave.

The specific value of $g(0)$ may seem arbitrary, because when no wealth is available the solution to the utility maximization problem is independent on the value of the focus weights. However, the value of $g(0)$ is important in determining how the salience of present and future consumption evolve as $b \to 0$. The condition $g(0) > 0$ implies that, when wealth approaches zero, present and future consumption become equally salient, independently on the shape of the functions $u_1(.)$, $u_2(.)$ and $f(.)$. Alternatively, if $g(0) = 0$, in some cases the ratio of the focusing weights $\frac{h_2\left(\frac{f(b_t)}{2}\right)}{h_1(b_t)}$ may diverge to infinity or go to zero. Hence, if $g(0) = 0$, as $b \to 0$ the present may become infinitely more salient than the future or vice versa. Condition $g(0) > 0$ rules out these situations.

Assumptions 1 and 2 will be maintained throughout the paper. In addition, I will often employ two additional assumptions, which are:

Assumption. *$g(x)$ is bounded above.*

Assumption (Functional forms)**.** *$f(x) = a \cdot x^{\alpha}$ for $\alpha \in (0,1)$ and $a > 0$; $u_1(x) = u_2(x) \equiv u(x) = \frac{(x+\epsilon)^{\sigma}}{\sigma}$ for $\epsilon \geq 0$ and $\sigma < 1$ (with the restriction $\epsilon > 0$ whenever $\sigma \leq 0$).*[5]

Assuming that $g(x)$ is bounded implies that $h(.)$ is always bounded. Boundedness implies that intensity of the utility of consumption decreases more rapidly than a logarithm for low utility levels, and less rapidly than a logarithm for high utility levels. Hence, $g(x)$ bounded above implies that the Weber-Fechner law holds in approximate terms, because it is always possible to approximate a logarithmic function with a bounded function. The second assumption (functional forms) is made

[5] This utility function is a HARA (Hyperbolic Absolute Risk Aversion) utility function. Because

$$-\frac{u''(x)}{u'(x)} = -\frac{\sigma - 1}{x + \epsilon}$$

the utility function displays constant relative risk aversion if $\epsilon = 0$; constant absolute risk aversion if both $\epsilon \to \infty$ and $\sigma \to -\infty$; decreasing absolute risk aversion otherwise. The parameter σ measures the curvature of the utility function. In particular, if $\sigma < 0$ the utility function is bounded above; if $0 \leq \sigma < 1$ the utility function is unbounded above.

for convenience. The use of any of the above two assumptions will be clearly stated when presenting the relevant propositions.

Under assumptions 1 and 2 the first-order condition of the utility maximization problem is:

$$u_1'(c_t) = \Delta(b_t) f'(b_t - c_t) u_2'(b_{t+1}) \tag{1}$$

where $\Delta(b_t)$ is the focus wedge:

$$\Delta(b_t) \equiv \frac{h_2\left(\frac{f(b_t)}{2}\right)}{h_1(b_t)}$$

The solution to the utility maximization problem is always unique for every b_t. Compared with the model described in the previous section, here initial wealth for generations $t > 1$ is determined endogenously and depends on the previous generation's savings decision. We argued in the previous section that the incentive to save may be increasing with initial wealth. The same argument implies here that bequests left may be a convex function of initial wealth.

The next three lemmas derive sufficient conditions under which $\Delta(b_t)$ is locally increasing for some b_t.

Lemma 3. *If $u_1'(0)$ is finite, the focus wedge is increasing in b_t for b_t sufficiently close to zero.*

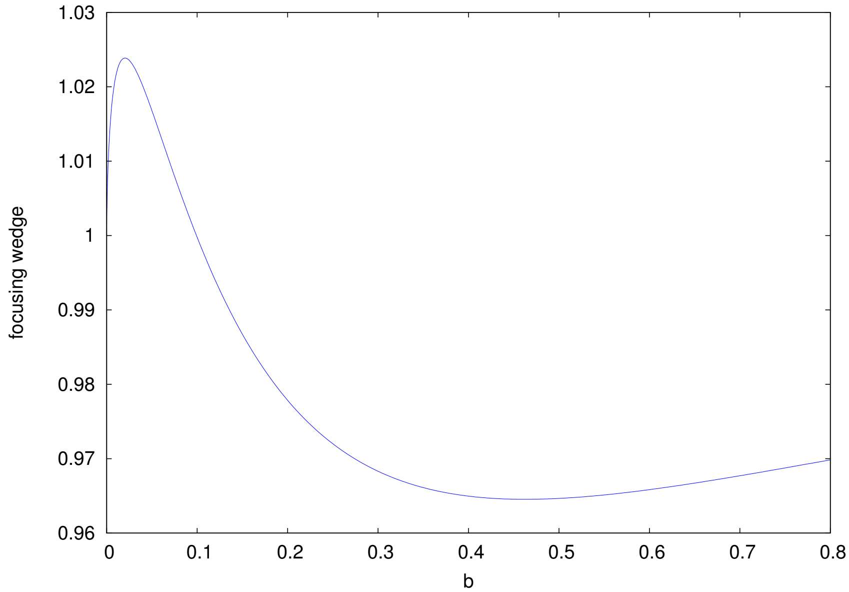
Proof. in appendix. □

When $u_1'(0)$, the utility of present consumption cannot grow arbitrarily fast. Therefore, for low b_t the driving force in the evolution of the focus wedge is the marginal return on investment. If the return on investment increases sufficiently fast with the size of the investment, then wealthier agents assign more weight to future consumption. Because the marginal return on investment is higher for low b_t, the focus wedge increases for low b_t. Finally, note that lemma 3 does not impose any restriction on $u_2'(0)$, meaning that the statement is true when $u_2'(0)$ is finite as well.

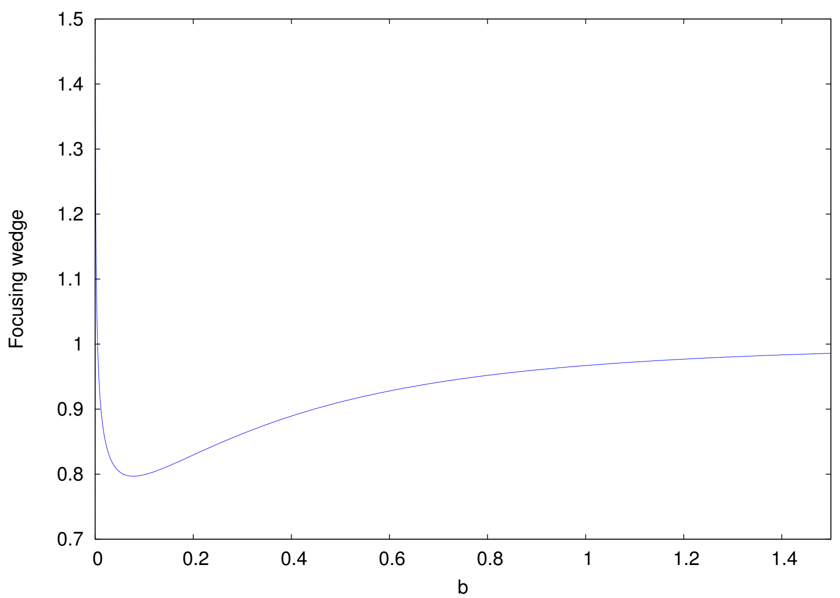
Lemma 4. *Assume that $u_2(x)$ is unbounded above, and $g(x)$ is bounded above. The focus wedge is somewhere increasing in b_t as long as $\exists b_t$ s.t. $u_2\left(\frac{f(b_t)}{2}\right) \neq u_1(b_t)$.*

Proof. in appendix. □

Boundedness of $g(x)$ implies that for very rich agents present and future consumption are approximately equally salient. When $g(0) > 0$, present and future are approximately equally salient also for very poor agents. Hence, unless the two focus weights are identical everywhere, the focus wedge must be increasing somewhere. Note that lemma 4 is true only if $u_2(x)$ is unbounded above. The reason is that if $u_2(x)$ is bounded above, the upper bound on $h_2(x)$ may be smaller than the upper bound on $h_1(x)$, implying that, by construction, for rich agents the future is less salient than the present. Finally, lemma 4 does not impose any restriction on $u_1(x)$, that could be unbounded as well.



(a) $u_1(x) = u_2(x) = \frac{1}{-5} \cdot (x+1)^{-5}$, $f(x) = x^{0.7}$ and $g(x) = (x+5)^{0.5}$, so that the conditions of lemma 3 and lemma 5 hold.



(b) $u_1(x) = u_2(x) = \frac{1}{0.7} \cdot x^{0.7}$, $f(x) = .5 \cdot x^{0.8}$ and $g(x) = 1 - e^{-5x}$, so that the conditions of lemma 4 hold.

Fig. 2: The focus wedge

Lemma 5. *Suppose that $u_1(x) = u_2(x) \equiv u(x)$ for all x, and that $u(x)$ is bounded above. The focus wedge is somewhere increasing in b.*

Proof. In appendix. □

The intuition of the above lemma is similar to the one described in lemma 4, because when the utility function is bounded also the salience of consumption is bounded.

3.1 The Steady State

To derive the steady state of the economy, I impose the functional form assumptions described in the previous section:

- $f(x) = a \cdot x^{\alpha}$ for $\alpha \in (0,1)$ and $a > 0$
- $u_1(x) = u_2(x) \equiv u(x) = \frac{(x+\epsilon)^{\sigma}}{\sigma}$ for $\epsilon \geq 0$ and $\sigma < 1$ (with the restriction $\epsilon > 0$ whenever $\sigma \leq 0$).

In steady state, $b_t = b_{t+1} = b_{ss}$ and $c_{ss} = b_{ss} - \left(\frac{2b_{ss}}{a}\right)^{\frac{1}{\alpha}}$, so that the steady-state level of bequests solves

$$\left(1 - \frac{\left(\frac{2b_{ss}}{a}\right)^{\frac{1}{\alpha}}}{b_{ss} + \epsilon}\right)^{\sigma - 1} = \alpha \cdot 2^{1-\frac{1}{\alpha}} \cdot a^{\frac{1}{\alpha}} \cdot \Delta(b_{ss})\,(b_{ss})^{\frac{\alpha-1}{\alpha}} \tag{2}$$

for

$$\Delta(b_{ss}) = \frac{h\left(\frac{a \cdot b_{ss}^{\alpha}}{2}\right)}{h(b_{ss})}$$

The left-hand side (LHS) of equation 2 is monotonically increasing in b_{ss}. On the right-hand side (RHS) of equation 2, $(b_{ss})^{\frac{\alpha-1}{\alpha}}$ is monotonically decreasing in b_{ss}, while $\Delta(b_{ss})$ is somewhere increasing. In other words, if $\Delta(b_{ss})$ were fixed, then LHS and RHS of equation 2 would cross only once and the model would have a unique steady state. However, because $\Delta(b_{ss})$ may be increasing for some b_{ss}, the RHS of equation 2 may be increasing for some b_{ss}, which implies that LHS and RHS of equation 2 may cross multiple times leading to multiple steady states.

Note the competing roles of the marginal return on investment and the focus wedge in determining the shape of the RHS of equation 2. The marginal return on investment always decreases with the size of the investment, making savings less appealing as wealth increases. At the same time, the focus wedge depends on the *total return* on investment and may increase with wealth, generating the opposite incentive. The relative importance of the marginal return on investment and the focus wedge in determining the shape of the RHS of equation 2 depends on the parameters α and a. If α is close to 1, then the marginal return on investment does not change much with the level of investment. If a is high enough, then the total return increases fast with the size of

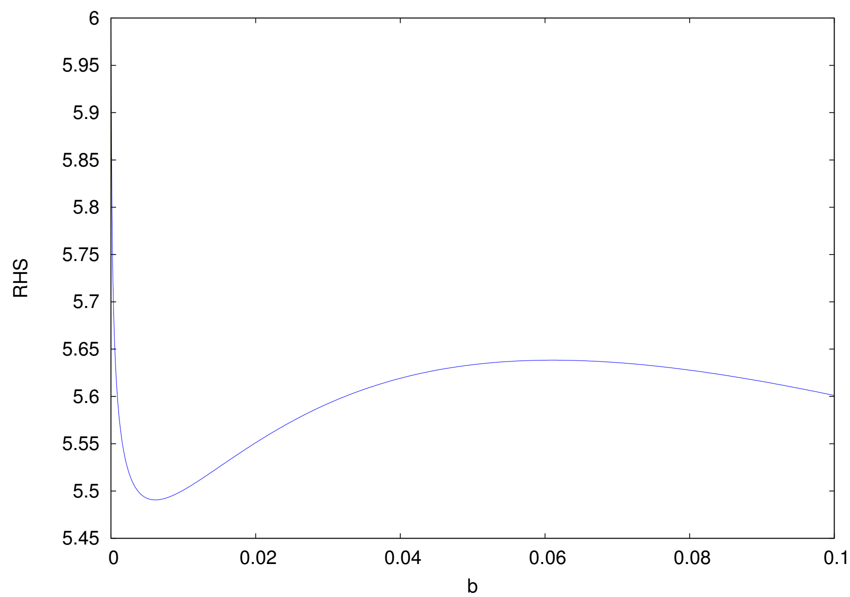


Fig. 3: RHS of equation 2 for $u(x) = \frac{1}{-5} \cdot (x+1)^{-5}$, $f(x) = 5 \cdot x^{0.98}$ and $g(x) = (x+5)^{0.5}$

the investment. If α is close to 1, and a is high, the shape of the RHS of equation 2 is mostly determined by the shape of the focus wedge.

Lemma 6. *If $a > 2$ and α arbitrarily close to 1 the RHS of equation 2 is increasing somewhere.*

Proof. In appendix. □

A necessary condition for the existence of a steady state is that the RHS of equation 2 is increasing somewhere. The above lemma shows that the production function, under some conditions, may cause a poverty trap. To better understand these condition, assume that the agent can invest in several projects, each of them with a given minimal and maximal scale. The agent will engage first in the projects with higher return, and later in projects with a lower return as the size of the investment increases. The resulting production function is $f(x)$. For the sake of the argument, assume that each of these projects has a linear return. Consider a specific high-return project, for example purchasing fertilizer. Lemma 6 shows that a poverty trap is possible if fertilizer is the best investment available for agents with different wealth levels. In this case, a poor agent who invests in fertilizer has a lower incentive to save than a richer agent who invests in fertilizer. The reason is that the salience of savings is greater for the agent who can invest more and rip a higher total return. The different initial incentives to save may translate in different steady states reached. In short, lemma 6 shows that a poverty trap can exist if both poor agents and less-poor agents have access to the same high marginal-return projects. On the other hand, there is no steady state

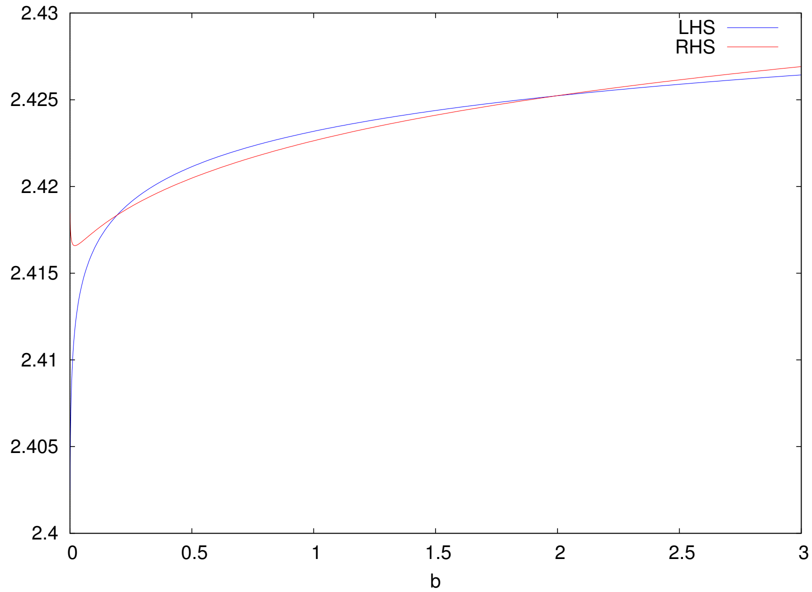


Fig. 4: LHS and RHS of equation 2 for $u(x) = \frac{1}{0.5} \cdot x^{0.5}$, $f(x) = 2.4103.x^{0.9995}$, $g(x) = (X+5)^{.2}$ (third steady state not shown)

multiplicity whenever different projects are pursued by agents with different wealth levels, because the marginal return on investment decreases with wealth, dampening the incentive to invest.

Proposition 7. *There is always at least one steady state. For $\alpha \to 1$, $\sigma > 0$, $\epsilon = 0$, there exist an $a > 2$ such that the economy has multiple steady states.*

Proof. In appendix. □

In the previous section, I argue that $\Delta(b)$ can be increasing because of technology, because of the curvature of the focus function, or because of the curvature of the utility function. In this section I show that the existence of a steady state is determined by the shape of the production function. A natural question arises: is it possible to find different sufficient conditions for the existence of multiple steady states, weaker than the ones of proposition 7, based on the shape of the utility function or the shape of the focus function?

To start with, note that if α is sufficiently low, then the marginal return on investment decreases rapidly with wealth. Regardless of the shape of the focus wedge, the economy has a unique steady state. It follows that multiple steady states can exist only if α is relatively high. In the limit case $\alpha \to 1$, one must assume that $a > 2$. The reason is that $c_{ss} = b_{ss} - \left(\frac{2b_{ss}}{a}\right)^{\frac{1}{\alpha}}$, meaning that if $\alpha \to 1$

and $a < 2$ then $c_{ss} \to 0$: consumption is zero in all steady states.[6] In other words, if $\alpha \to 1$, then $a > 2$ is a necessary condition for the existence of steady states with positive consumption.

Of course, the previous discussion does not imply that $\alpha \to 1$ and $a > 2$ are necessary conditions for steady-state multiplicity. It may be possible to find different necessary conditions, where α is high but bounded away from 1, and multiple steady states exist because of the curvature of the utility function or the curvature of the focus function. I do not explore this possibility here. However, in section 5 I assume that there is a perfect credit market, so that the return on savings is linear. I show that a poverty trap emerges whenever the utility function is bounded above.

4 Dynamics

In section 3, I showed that the solution to the consumption-savings problem is unique for every level of initial assets. Call $b^{\star}_{t+1}(b_t)$ the amount of bequests left as a function of the amount of bequests received. Under the same functional form assumptions made in the previous section, $b^{\star}_{t+1}(b_t)$ is implicitly defined as:

$$b^{\star}_{t+1}(b_t) \equiv b_{t+1} : \left(b_t - \left(\frac{2}{a} b_{t+1} \right)^{\frac{1}{\alpha}} + \epsilon \right)^{\sigma-1} = a\alpha\Delta(b_t) \left(\frac{2}{a} b_{t+1} \right)^{\frac{\alpha-1}{\alpha}} (b_{t+1} + \epsilon)^{\sigma-1} \tag{3}$$

where

$$\Delta(b_t) = \frac{h_2 \left(\frac{a}{2} b_t^{\alpha} \right)}{h_1 (b_t)}$$

The LHS of equation 3 is increasing in b_{t+1} and decreasing in b_t. The RHS of 3 is decreasing in b_{t+1} and depends on b_t only through the focus wedge. It follows that if the focus wedge is constant or increasing, $b^{\star}_{t+1}(b_t)$ is always increasing. However, if $\Delta(b_t)$ is somewhere decreasing, $b^{\star}_{t+1}(b_t)$ may also be decreasing. Intuitively, if the salience of future consumption decreases in wealth over some range, then wealthier agents may save less than poorer agents do.

Lemma 8. *If $g(x)$ is bounded above, there exist a σ arbitrarily close to 1 such that $b'_{t+1}(b_t) < 0$ for some b_t.*

Proof. In appendix. □

When the salience of consumption is bounded above, then the focus wedge cannot be monotonically increasing everywhere. For low wealth levels, the return on savings is high (and increasing with wealth) and the salience of future consumption is greater than the salience of present consumption, so that $\Delta(b_t) > 1$. But when $g(x)$ is bounded above, for b_t large the salience of present consumption approaches the salience of future consumption and $\Delta(b_t)$ converges to 1.

[6] It is possible to have steady states where consumption is zero but bequests are positive. However, I ruled them out by assuming that $\epsilon = 0$ and $\sigma \in (0, 1)$, so that the marginal utility of consumption is infinity at zero.

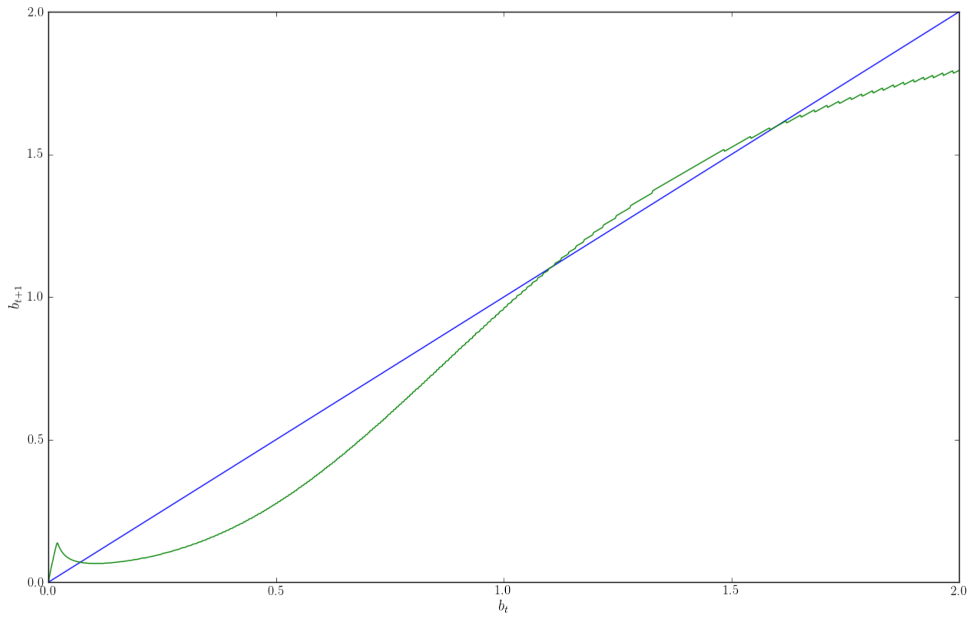


Fig. 5: Non-monotonic savings function.

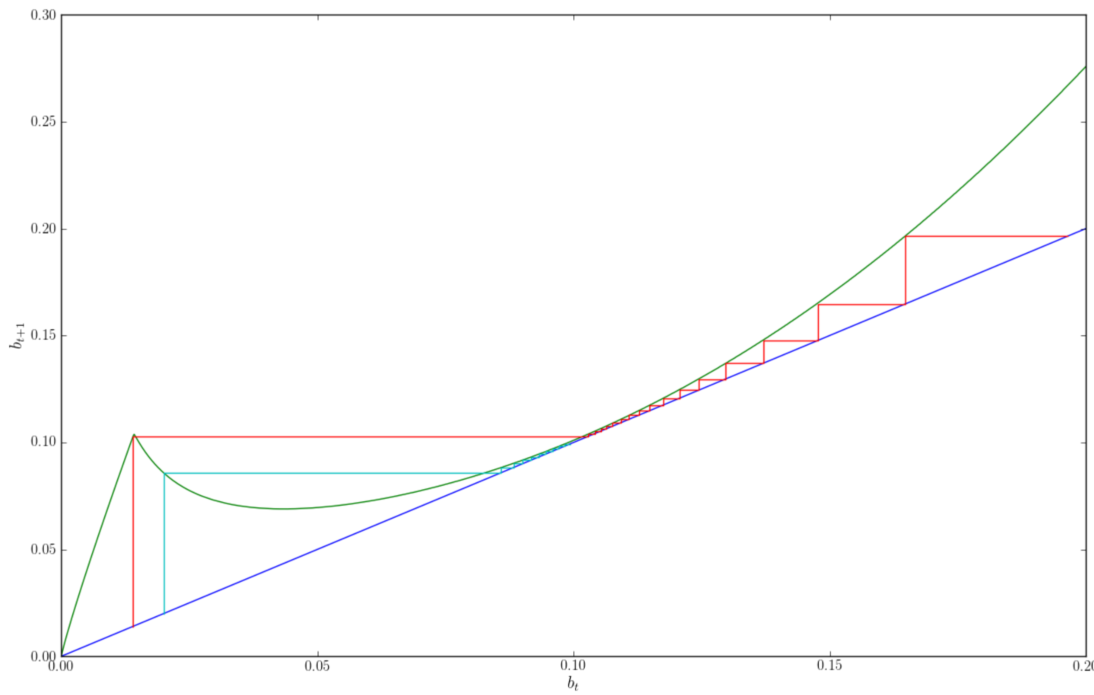


Fig. 6: Non monotonic convergence to the steady state (highest steady state not shown).

It follows that the convergence to the steady state may be non-monotonic, meaning that a lower starting condition may lead to the convergence to a higher steady state. Hence, next to poverty traps, the economy may display *middle-income traps.* Figures 5 and 6 illustrate this possibility. For low and high starting conditions the household converges to the high steady state. However, for intermediate starting conditions, the household converges to a low steady state. Intuitively, poor households save a large fraction of their wealth because the return on saving is high and very salient. Rich households save a low fraction of their wealth, but if the household is sufficiently rich, the amount of bequests left is enough to reach the high steady state. Households with intermediate wealth do not save enough and remain stuck in the low steady state.

5 Perfect Credit Market

So far I assume that agents cannot borrow or lend, but only invest in their own production function, and I showed that multiple steady states are possible. However, the reader may wonder whether the poverty trap arises because of the focusing effect, or because of the interaction between the focusing effect and the absence of a credit market. To address this question, here I introduce a perfect credit market into the model. I show that, under a boundedness assumption on the utility function, multiple steady states are possible here as well.

Assume that agents can borrow and lend at an interest rate r. For any technology $f(x)$ satisfying the standard Inada conditions, a perfect credit market implies two things:

1. In every period, the agent can borrow at rate r and invest in $f(x)$ until $f'(x) = 1 + r$. This is equivalent to assuming that the agent receives a lump-sum payment $y(r)$, increasing in r and equal to the infra-marginal benefit of borrowing at rate r and investing in $f(x)$.

2. After receiving $y(r)$, the agent saves linearly at the interest rate r.

The consumption-savings problem is now

$$\max_{c_t, b_{t+1}} \left\{ h_1 \left(b_t + y(r)\right) u_1(c_t) + h_2 \left(\frac{\left(b_t + y(r)\right)(1+r)}{2} \right) 2 \cdot u_2(b_{t+1}) \right\}$$

$$s.t \quad \left(b_t + y(r) - c_t\right)(1+r) = 2b_{t+1}$$

When analyzing the case of no credit market, I assumed that the technology is close to linear with slope $a > 2$, implying a return on savings above 100% for *some* saving levels. With a perfect credit market, the return on savings is exactly linear and the logic behind proposition 7 applies here as well: it is possible to show that for some $r > 1$ a poverty trap exists. However, assuming a return on savings above 100% for *any* saving levels is quite unreasonable. I will therefore limit the analysis to the case $r < 1$.

In a steady state, $b_t = b_{t+1} = b_{ss}$ and $c_{ss} = y(r) - \frac{b_{ss}(1-r)}{(1+r)}$. Assuming again that $u_1(x) = u_2(x) = \frac{1}{\sigma}(x+\epsilon)^\sigma$ for $\sigma < 1$ and $\epsilon \geq 0$ (with $\epsilon > 0$ if $\sigma \leq 0$), the steady-state level of bequests solves

$$\left(\frac{y(r) - b_{ss}\frac{(1-r)}{(1+r)} + \epsilon}{b_{ss} + \epsilon} \right)^{\sigma - 1} = (1+r)\Delta(b_{ss}) \tag{4}$$

The LHS of equation 4 is increasing in b_{ss}. The shape of the RHS instead depends on the shape of $\Delta(b_{ss})$.

Lemma 9. *Assume that $r < 1$, and that $\sigma < 0$ (so that the utility function is bounded above). $\Delta(b_{ss})$ is increasing somewhere.*

Proof. In appendix. □

When there is no credit market, the present is more or less salient than the future depending on the return on investment at a specific b. In particular, for b small the future will be more salient than the present, while the opposite is true for large b. Instead, with perfect credit market the salience of future consumption relative to present consumption depends on r . If $r < 1$ the future is always less salient than the present. However, if the utility function is bounded above, then the difference in salience between present and future becomes smaller as b_{ss} increases. The future is discounted less and less, meaning that wealthier agents save a larger fraction of their initial wealth compared to poorer agents.[7]

Figure 7 illustrates a numerical examples, in which under the conditions assumed in lemma 9 multiple steady states exist. Finally, it is also possible to show that multiple steady states emerge when the utility function is unbounded but the function $g()$ is bounded above. The reason is that, also in this case, the salience of consumption is bounded above and the distortion introduced by the focusing effect decreases as wealth increases.

6 Commitment Savings

When preferences are distorted by the focusing effect, agents' choices depend on the set of available choices in a way that may lead to a poverty trap. It is therefore interesting to note that, in a consumption-savings set up, people often strategically manipulate their choice set by mean of various commitment-saving devices. The typical explanation for the use of commitment devices relies on time inconsistency. In this section, I show that commitment increases savings also when the relevant behavioral bias is the focusing effect, in a way that empirically distinguishable from a situation in which the only behavioral bias is time inconsistency.

[7] Bounded utility functions have been frequently reported in the literature. See, for example, Havránek, Horvath, Iršová, and Rusnak (2013) for a meta-analysis or the cross-country estimates. The utility function used here does not display constant elasticity of substitution (CES), but can approximate a CES function by choosing ϵ small.

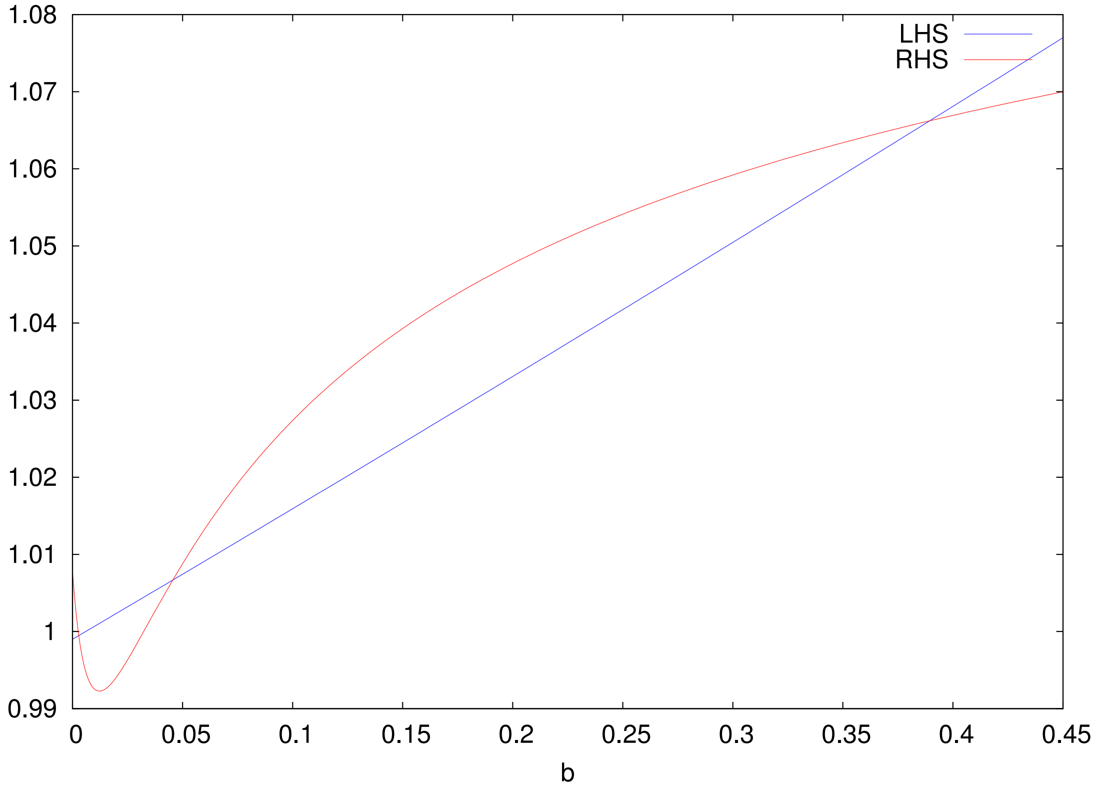


Fig. 7: LHS and RHS of equation 4 for $g(x) = \log(x/0.0005 + 2)$, $u(x) = -\frac{1}{0.01}(x + 10)^{-0.01}$, $r = .2$, $y(r) = 0.01$.

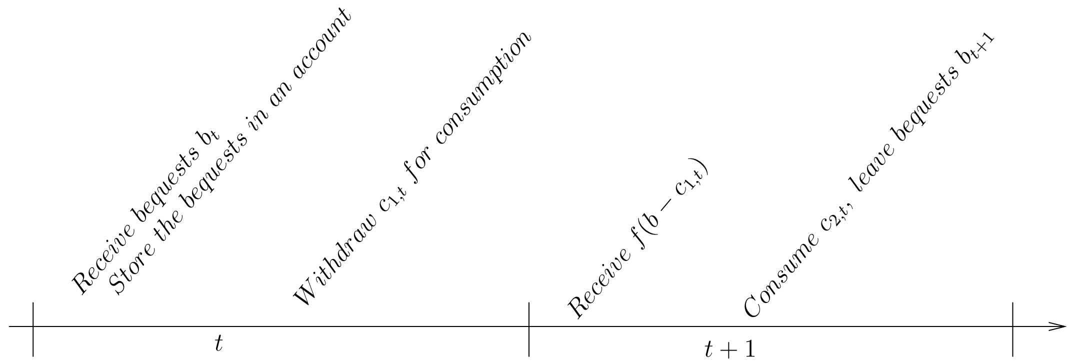


Fig. 8: Timeline with choice of accounts.

Assume that the first period of an agent's life is divided into two sub-periods. In the first sub-period, an agent receives her bequests and stores them into an account. In the second sub-period, she withdraws some money for consumption, and saves the rest. There are two types of accounts, a normal account for simple wealth storage, and a commitment account that is subject to a withdrawal tax if the account balance drops below a given threshold during period 1.

From the agent's point of view, adopting the commitment account has one obvious implication: if the threshold that triggers the punishment is above what the agent would have saved without commitment, then adopting the commitment account may cause the agent to increasing her savings in order to avoid the punishment. However, in this context, the commitment account increases savings through a second channel. Remember that the focus weight on present consumption is a function of the utility that can be achieved today if savings are set to zero. If the agent adopts the normal account, the focus weight on present consumption is $h(b_t)$. If the agent adopts the commitment account, the focus weight on present consumption is $h(b_t - \tau\kappa)$ where τ is the tax and κ is the threshold (the focus weight on future consumption is unchanged). In other words, the commitment mechanism affects the choice set and decreases the maximum utility level that is achievable today, which in turns decreases the salience of present consumption relative to future consumption.

It follows that the focusing effect has one distinctive empirical implication. Without the focusing effect, whenever an agent anticipates that she will not save enough she may adopt a commitment account similar to the one previously discussed. The commitment savings account may push the agent from saving below the threshold to saving exactly at the threshold, but never above the threshold. The focusing effect introduces an additional element: the commitment threshold makes consumption less salient. This implies that some people may save below the threshold, adopt the commitment technology, and start saving above the threshold. Alternatively, they may already save above the threshold, but nonetheless adopt the commitment device to save even more.

When comparing the distribution of savings between a treatment group to which the commitment account was offered and a control group, the presence (or absence) of the focusing effect can be detected by comparing the two distributions of savings. Without the focusing effect, offering the account causes all saving levels below κ to (weakly) loose mass, causes the savings level κ to gain mass, and causes no changes in the proportion of agents saving strictly above κ. The presence of the focusing effect instead implies that savings levels above κ may also gain mass.

Why would the agent choose a commitment-savings account? A natural normative benchmark is a situation in which, at the beginning of life, the agent has full control over her lifetime consumption path. In this benchmark, when young the agent can decide that in period 2 she will consume all her wealth, or that she will leave all her wealth as bequests. The benchmark focus wedge is

$$\Delta^\star(b) \equiv \frac{h(f(b))}{h(b)} > \Delta(b) \equiv \frac{h\left(\frac{f(b)}{2}\right)}{h(b)}$$

which is always greater than the focus wedge used by the agent when deciding how much to save. In other words, in this normative benchmark the agent places more weight on the future than in equilibrium. If the agent uses $\Delta^\star(b)$ to discount the future when the type of account is chosen, then some level of commitment is valuable for every wealth level.[8] By adopting the commitment account, the agent can make the present less salient, and therefore push the focus wedge toward $\Delta^\star(b)$.

Lemma 10. *Suppose that a commitment account can be purchased at a cost and that, once the account is purchased, the agent can set the punishment $\tau\kappa$ optimally. There exists a $\underline{b}$ such that all agents with wealth level below $\underline{b}$ do not purchase the account. Furthermore, if the focus function $g(x)$ is bounded above, there is also a $\bar{b}$ such that all agents with wealth level above $\bar{b}$ do not purchase the account.*

Proof. In appendix. □

When $g(x)$ is bounded above, the value of commitment goes to zero for $b \to \infty$ and for $b \to 0$. If the commitment device can be adopted at a cost, then the poorest and the richest agents do not purchase commitment, but agents with intermediate wealth levels might purchase it. A similar result holds if the account is free, but the punishment $\tau\kappa$ is given: very rich and very poor agent do not adopt the account, but other agents might.

[8] A second possibility is that multiple behavioral biases are at play. The agent chooses the commitment account because she anticipates to be dynamically inconsistent. Once the account is chosen, the focusing effect determines how savings respond to the presence of the punishment.

7 Conclusions

I develop a consumption-savings model where agents' choices are distorted by the focusing effect: when choosing from a choice set, a decision maker overweights the goods in which her options differ the most. It follows that, as wealth increases, the salience of consumption today relative to consumption tomorrow changes. In particular, if the marginal return on investment is sufficiently high, then the salience of future consumption relative to present consumption increases with wealth because, as wealth grows, future consumption possibilities expand faster than present consumption possibilities.

I show that, if the marginal return on investment is high and flat, then a poverty trap may emerge. In this case, the *percentage* return on investment at different wealth levels is approximately constant, but the *total* return on investment increases rapidly with wealth. Because the salience of future consumption depends on the total return, wealthier agents place more importance on future consumption and save more than poorer agents do. Wealth inequality and poverty are transmitted from generation to generation and a poverty trap may emerge.

Next to poverty traps, middle-income traps are possible. The salience of future consumption relative to present consumption may increase with wealth over some wealth levels, but decrease with wealth over some other wealth levels. It follows that some poor agents may save more than richer agents. When this happens, the convergence to different steady states may be non-monotonic in the initial wealth level: the steady-state wealth level reached by households starting with low wealth may be higher than the steady-state wealth level reached by households starting with higher wealth. Whereas poor and rich households converge to the same steady state, middle income households are stuck in a low steady state. I show that a middle income trap emerges when the salience of consumption is bounded above. For low wealth levels, the marginal return on savings is very high, which implies that the relative salience of future consumption is high and increasing with wealth. But for high wealth levels, boundedness of salience implies that the focusing effect is not relevant in the consumption-savings decision, and that the incentive to save is relatively low. If starting wealth is sufficiently large, then the households will nonetheless converge to the high steady state. But if the starting wealth is in some intermediate range, the household will converge to a low steady state.

I also consider the case of a perfect credit market. If the utility function is bounded above, then a poverty trap may exist also in this case. In utility terms, the difference between consuming in the first period of life and in the second period of life becomes smaller as wealth increases. Therefore, the distortion introduced by the focusing effect becomes less severe as wealth increases, so that rich agents have higher incentives to save than poor agents.

Finally, I argue that when the preferences are distorted by the focusing effect, commitment-saving devices increase savings in a way that is empirically distinct from other behavioral biases.

With the focusing effect, when a punishment is imposed on the agent for dropping savings below a given threshold, the agent increases the level of savings even when this threshold is not binding. The reason is that the punishment reduces the maximum utility achievable in the current period of life and the salience of present consumption. Measuring the empirical relevance of the focusing effect relative to other behavioral bias in distorting savings decisions is left for future work.

A Appendix: A Rational Inattention Interpretation

The consumption-savings problem I analyzed in this paper can also be interpreted as the reduced form of a rational-inattention model. Assume that, in every period, after the agent decides on the amount to save and to consume two types of mistakes can occur. In one type of mistake, consumption is destroyed but savings are left untouched. In the other type of mistake savings are destroyed but consumption is left untouched. I assume that the probability of each mistake is independent on the realization of the other mistake.

The agent can monitor the two consumption sets in search of potential mistakes and correct them. However, monitoring requires effort, and effort is costly. Call p_c the probability of a mistake in which consumption is destroyed, and p_s the probability of a mistake in which savings are destroyed. I assume that

$$1 - p_c = e_c$$

$$1 - p_s = e_s$$

The cost of monitoring depends on the size of the two choice sets (present and future) available to the agent, and is assumed:

$$C(e_c + e_s) = \mu_c\left(b, f(b)\right)\frac{e_c^2}{2} + \mu_s\left(b, f(b)\right)\frac{e_s^2}{2}$$

where $\mu_c\left(b, f(b)\right)$ and $\mu_s\left(b, f(b)\right)$ are increasing in both arguments. The optimal-effort problem is

$$\max_{e_c, e_s}\left\{e_c u_1(c_t) + 2e_s u_2(b_{t+1}) - \left[\mu_c\left(b, f(b)\right)\frac{e_c^2}{2} + \mu_s\left(b, f(b)\right)\frac{e_s^2}{2}\right]\right\}$$

with solution

$$e_c^\star = \frac{u_1(c)}{\mu_c\left(b, f(b)\right)}$$

$$e_s^\star = \frac{u_2\left(b_{t+1}\right)}{2 \cdot \mu_s\left(b, f(b)\right)}$$

I make the following assumption

Assumption 11.

$$\frac{\mu_s\left(b, f(b)\right)}{\mu_c\left(b, f(b)\right)} = \frac{\tilde{h}_c(b)}{\tilde{h}_s(f(b))}$$

where $\tilde{h}_c(x)$ and $\tilde{h}_s(x)$ are increasing functions.

Under the above assumption, the optimal $e_c^\star/e_s^\star$ increases with the size of the current consumption choice, and decreases with the size of the future consumption choice. Intuitively, if one of the two choice sets increases, the impact of mistakes is minimized by monitoring relatively more the choice set that increased the most. Under this assumption, this costly-attention model replicates the main feature of the model discussed in the body of the paper: as wealth increases, the agent changes her relative valuation of future consumption as a function of how rapidly the future consumption possibilities expand relative to present consumption possibilities.

By assuming that $u_1(x) = u_2(x) = x^\sigma$ for $\sigma \in (0, \frac{1}{2})$, the consumption-savings problem becomes:

$$\frac{\mu_c\left(b, f(b)\right)}{\tilde{h}_c(b)} \cdot \max_{c_t, b_{t+1}} \left\{ \tilde{h}_c(b) \cdot c_1^{2\sigma} + \tilde{h}_s(f(b)) b_{t+1}^{2\sigma} - \left[\mu_c\left(b, f(b)\right) \frac{(e_c^\star)^2}{2} + \mu_s\left(b, f(b)\right) \frac{(e_s^\star)^2}{2} \right] \right\}$$

$$s.t \;\; f(b_t - c_1) = 2b_{t+1}$$

which is almost equivalent to the consumption-savings problem discussed previously. The two problems become identical by setting

$$\tilde{h}_c(b_t) = g(b_t^{2\sigma});\; \tilde{h}_s(f(b_t)) = 2 \cdot g\left(\left(\frac{f(b_t)}{2} \right)^{2\sigma} \right)$$

B Appendix: Mathematical Derivations

Proof of lemma 3.

Because $g(0) > 0$, $\Delta(0) = \frac{g(0)}{g(0)} = 1$. On the other hand, we have

$$\lim_{b\to 0} \left[\frac{u_2\left(\frac{f(b)}{2}\right) - u_2(0)}{u_1\left(b\right) - u_1(0)} \right] = \lim_{b\to 0} \left[\frac{u_2'\left(\frac{f(b)}{2}\right) \frac{f'(b)}{2}}{u_1'\left(b\right)} \right] = \infty$$

because, $\lim_{b\to 0} \left\{ \frac{f'(b)}{2} \right\} = \infty$ and $u_2'(0)$ is either a positive number or diverges to infinity as well. It follows that, for b sufficiently small

$$u_1\left(b\right) - u_1(0) < u_2\left(\frac{f(b)}{2}\right) - u_2(0)$$

and

$$\Delta(b) = \frac{g(\frac{f(b)}{2})}{g(b)} > 1$$

therefore $\Delta(b)$ must be increasing for b small.

Proof of lemma 4.

Because $g(0) > 0$, $\Delta(0) = 1$. In addition, $\lim_{b\to\infty}\Delta(b) = \lim_{b\to\infty}\left(\frac{g\left(u_2\left(\frac{f(b)}{2}\right)-u_2(0)\right)}{g(u_1(b)-u_1(0))}\right) \geq 1$, with equality if $u_1(x)$ is unbounded above, and strict inequality if $u_1(x)$ is bounded above. Because $\Delta(b)$ is not everywhere identical to one, the focus wedge must be increasing somewhere.

Proof of lemma 5.

We have $\Delta(0) = 1$, and $\lim_{b\to\infty}\left(\frac{g\left(u\left(\frac{f(b)}{2}\right)-u(0)\right)}{g(u(b)-u(0))}\right) = 1$. Because $\Delta(b)$ is not everywhere identical to one, the focus wedge must be increasing somewhere.

Proof of lemma 6.

Define

$$\lim_{\alpha\to1} a^{\frac{1}{\alpha}}\cdot 2^{1-\frac{1}{\alpha}}\alpha\frac{g\left(\frac{1}{\sigma}\left(\left(\frac{a\cdot b_{ss}^{\alpha}}{2}+\epsilon\right)^{\sigma}-\epsilon^{\sigma}\right)\right)}{g\left(\frac{1}{\sigma}\left((b_{ss}+\epsilon)^{\sigma}-\epsilon^{\sigma}\right)\right)} = a\cdot\frac{g\left(\frac{1}{\sigma}\left(\left(\frac{a\cdot b_{ss}}{2}+\epsilon\right)^{\sigma}-\epsilon^{\sigma}\right)\right)}{g\left(\frac{1}{\sigma}\left((b_{ss}+\epsilon)^{\sigma}-\epsilon^{\sigma}\right)\right)} \equiv \kappa(b)$$

we know that $\kappa(0) = a$. In addition, if $a > 2$, $\frac{1}{\sigma}\left(\left(\frac{a\cdot b_{ss}}{2}+\epsilon\right)^{\sigma}-\epsilon^{\sigma}\right) > \frac{1}{\sigma}\left((b_{ss}+\epsilon)^{\sigma}-\epsilon^{\sigma}\right)$ for all b_{ss}. Hence $\kappa(b_{ss}) > a$ for all $b_{ss} > 0$ meaning that $\kappa(b_{ss})$ is somewhere increasing. Finally, note that the RHS of equation 2 tends to infinity for $b_{ss}\to0$, and to zero for $b_{ss}\to\infty$. At the same time, for every $b_{ss} > 0$, it is possible to find an α arbitrarily close to one, such that the distance between the RHS of equation 2 and $\kappa(b_{ss})$ is arbitrarily small. Hence, for every $b_{ss} > 0$ such that $\kappa(b_{ss})$ is increasing, it is possible to find an α sufficiently large such that the RHS of equation 2 is also increasing.

Proof of proposition 7.

Simple algebra shows that the RHS of equation 2:

$$\left(1-\frac{\left(\frac{2b_{ss}}{a}\right)^{\frac{1}{\alpha}}}{b_{ss}+\epsilon}\right)^{\sigma-1} \tag{5}$$

is strictly increasing for $b_{ss}\in\left(0,\left(\frac{a}{2}\right)^{\frac{1}{1-\alpha}}\right)$. The LHS of equation 2

$$2^{1-\frac{1}{\alpha}}a\Delta(b_{ss})\alpha\left(b_{ss}\right)^{\frac{\alpha-1}{\alpha}} = 2^{1-\frac{1}{\alpha}}a^{\frac{1}{\alpha}}\frac{g\left(\frac{1}{\sigma}\left(\left(\frac{a\cdot b_{ss}^{\alpha}}{2}+\epsilon\right)^{\sigma}-\epsilon^{\sigma}\right)\right)}{g\left(\frac{1}{\sigma}\left((b_{ss}+\epsilon)^{\sigma}-\epsilon^{\sigma}\right)\right)}\alpha b_{ss}^{1-\frac{1}{\alpha}} \tag{6}$$

is going to zero for $b_{ss} \to \infty$, and to infinity for $b_{ss} \to 0$. Hence equation 2 has at least one solution in $\left(0, \left(\frac{a}{2}\right)^{\frac{1}{1-\alpha}}\right)$.

Because $a > 2$, if $\epsilon = 0$ as $\alpha \to 1$ expression 5 becomes a straight line at $\left(1 - \frac{2}{a}\right)^{\sigma-1}$. In addition, if $a > 2$ and $\alpha \to 1$ expression 6 is arbitrarily close to $a\Delta(b_{ss})$ for all $b_{ss} > 0$, where

$$a\Delta(b_{ss}) = a\frac{g\left(\frac{1}{\sigma}\left(\frac{a \cdot b_{ss}}{2}\right)^{\sigma}\right)}{g\left(\frac{1}{\sigma}\left(b_{ss}\right)^{\sigma}\right)}$$

starts at a and is always above a. If $a < \left(1 - \frac{2}{a}\right)^{\sigma-1}$ but $a \approx \left(1 - \frac{2}{a}\right)^{\sigma-1}$ equation 2 has three solutions: at $b_{ss} = 0$ the LHS of 2 diverges to infinity while the RHS of 2 is finite; for $b_{ss} > 0$ but arbitrarily small the LHS of 2 is approximately equal to $a\Delta(b_{ss})$, which is approximately equal to a and is below the RHS of 2; for $b_{ss} > 0$ larger the LHS of 2 is approximately equal to $a\Delta(b_{ss})$, which is above a and is above the RHS of 2; for b_{ss} sufficiently large the LHS of 2 goes to zero while the RHS of 2 is positive.

Proof of lemma 8.

Using the implicit function theorem, it is possible to show that the sign of $b'_{t+1}(b_t)$ is equal to the sign of

$$\frac{\partial\left[\Delta(b_t)\right]}{\partial b_t}\left(a\alpha\left(\frac{2}{a}b^{\star}_{t+1}(b_t)\right)^{\frac{\alpha-1}{\alpha}}\left(b^{\star}_{t+1}(b_t)\right)^{\sigma-1}\right) - (\sigma-1)\left(b_t - \left(\frac{2}{a}b^{\star}_{t+1}(b_t)\right)^{\frac{1}{\alpha}}\right)^{\sigma-2} \tag{7}$$

for $\sigma \to 1$ the above expression becomes

$$\frac{\partial\left[\Delta(b_t)\right]}{\partial b_t} \cdot a\alpha\left(\frac{2}{a}b^{\star}_{t+1}(b_t)\right)^{\frac{\alpha-1}{\alpha}}$$

and its sign depends on the sign of $\frac{\partial[\Delta(b_t)]}{\partial b_t}$. Assuming that $g(0) > 0$, implies that $\Delta(0) = 1$. Furthermore, when $g()$ is bounded above

$$\lim_{b\to\infty}\Delta(b) = \lim_{b\to\infty}\left[\frac{g\left(\frac{1}{\sigma}\left(\frac{a}{2}b^{\alpha}+\epsilon\right)^{\sigma} - \frac{1}{\sigma}\left(\epsilon\right)^{\sigma}\right)}{g\left(\frac{1}{\sigma}\left(b+\epsilon\right)^{\sigma} - \frac{1}{\sigma}\left(\epsilon\right)^{\sigma}\right)}\right] = \begin{cases}\lim_{b\to\infty}\left[\frac{g\left(-\frac{1}{\sigma}(\epsilon)^{\sigma}\right)}{g\left(-\frac{1}{\sigma}(\epsilon)^{\sigma}\right)}\right] & \text{if } \sigma < 0\\ \lim_{x\to\infty}\left[\frac{g(x)}{g(x)}\right] & \text{if } \sigma \geq 0\end{cases} = 1$$

Because $\Delta(b) \neq 1$ for some $b > 0$, $\Delta(b)$ must be decreasing somewhere.

Proof of lemma 9.

Write

$$\Delta(b_{ss}) = \frac{g\left(\frac{1}{\sigma}\left(\frac{(b_{ss}+y(r))(1+r)}{2}+\epsilon\right)^{\sigma} - \frac{1}{\sigma}(\epsilon)^{\sigma}\right)}{g\left(\frac{1}{\sigma}\left(b_{ss}+y(r)+\epsilon\right)^{\sigma} - \frac{1}{\sigma}(\epsilon)^{\sigma}\right)}$$

which is always below one if $r < 1$. Also, if $\sigma < 0$ utility is bounded, and both numerator and denominator converge to $g(-\frac{1}{\sigma}(\epsilon)^{\sigma})$ as $b_{ss} \to \infty$. It follows that $\lim_{b\to\infty} \Delta(b_{ss}) = 1$. Hence $\Delta(b_{ss})$ must be increasing somewhere.

Proof of lemma 10.

It follows simply because $\lim_{b\to 0}\left[\frac{h\left(\frac{f(b)}{2}\right)}{h(b)}\right] = \lim_{b\to 0}\left[\frac{h(f(b)}{h(b)}\right] = 1$, and because because $\lim_{b\to\infty}\left[\frac{h\left(\frac{f(b)}{2}\right)}{h(b)}\right] = \lim_{b\to\infty}\left[\frac{h(f(b)}{h(b)}\right] = 1$.

References

Ananth, B., D. Karlan, and S. Mullainathan (2007). Microentrepreneurs and their money: Three anomalies. *Working paper*.

Banerjee, A. and E. Duflo (2011). *Poor economics: a radical rethinking of the way to fight global poverty*. PublicAffairs.

Banerjee, A. and S. Mullainathan (2010). The shape of temptation: Implications for the economic lives of the poor. *working paper*.

Banerjee, A. V. and S. Mullainathan (2008). Limited attention and income distribution. *The American Economic Review 98*(2), 489–493.

Banerjee, A. V. and A. F. Newman (1993). Occupational choice and the process of development. *Journal of political economy*, 274–298.

Beaman, L., J. Magruder, and J. Robinson (2013). Minding small change: Limited attention among small firms in kenya. *Working paper*.

Bernheim, B. D., D. Ray, and S. Yeltekin (2013, January). Poverty and self-control. Working Paper 18742, National Bureau of Economic Research.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). Salience and consumer choice. *Journal of Political Economy 121*(5), 803–843.

Duflo, E., M. Kremer, and J. Robinson (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *The American Economic Review 101*(6), 2350–2390.

Gabaix, X. (2012). A sparsity-based model of bounded rationality. *working paper*.

Galor, O. and J. Zeira (1993). Income distribution and macroeconomics. *The review of economic studies 60*(1), 35–52.

Havránek, T., R. Horvath, Z. Iršová, and M. Rusnak (2013). Cross-country heterogeneity in intertemporal substitution. *Working paper*.

Kahneman, D., A. B. Krueger, D. Schkade, N. Schwarz, and A. A. Stone (2006). Would you be happier if you were richer? a focusing illusion. *Science 312*(5782), 1908–1910.

Kőszegi, B. and A. Szeidl (2013). A model of focusing in economic choice. *The Quarterly Journal of Economics 128*(1), 53–104.

Laibson, D., A. Repetto, and J. Tobacman (2003). A debt puzzle. In S. J. W. M. Aghion P, Frydman R (Ed.), *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, pp. 228. Princeton University Press.

Moav, O. (2002). Income distribution and macroeconomics: the persistence of inequality in a convex technology framework. *Economics Letters 75*(2), 187–192.

Moav, O. and Z. Neeman (2010). Status and poverty. *Journal of the European Economic Association 8*(2-3), 413–420.

Moav, O. and Z. Neeman (2012). Saving rates and poverty: The role of conspicuous consumption and human capital. *Economic Journal 122*(563), 933–956.

Schkade, D. A. and D. Kahneman (1998). Does living in california make people happy? a focusing illusion in judgments of life satisfaction. *Psychological Science 9*(5), 340–346.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics 50*(3), 665–690.

Udry, C. and S. Anagol (2006). The return to capital in ghana. *American Economic Review 96*(2), 388–393.