Helsinki Center of Economic Research

Discussion Papers

Existence Effects of Retirement Savings Program

Timo Seppälä
Centre for Health Economics at STAKES, RUESG and HECER

Discussion Paper No. 142
January 2007

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781, E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 142

Existence Effects of Retirement Savings Program*

Abstract

There is a growing tendency to put more pressure on individual to take care of own pension. Current paper concentrates on existence effects of a lucrative retirement savings program that pays better interest rate on savings than saving in checking account but enrolling on it incurs a utility cost. Study objects are three hyperbolic agent types: naive, sophisticate and learning naive, where the first two are well known types while the third is a new type with a learning capability. The agent plans simultaneously intertemporal consumption and when to enrol (if ever) on the savings program. The main results establish that for the same range of the effort cost naive suffers from, sophisticate never suffers from and always gains from and learning naive can suffer or gain from the existence of the saving program in terms of retirement savings. If the effort cost is at low or at intermediate level learning naive gains from the existence of the savings program but starts to suffer from it when the effort cost level increases to relatively high or very high level. Finally, if lowering the effort cost is expensive relative to hastening the learning capability, possibilities to learn from self-control problems should be improved. If those both are expensive, it should be tried to avoid heterogeneity in different possibilities to save or let the effort cost to rise so that only sophisticates start the savings program while naives and learning naives abandon it fairly quickly.

JEL Classification: D19, D91.

Keywords: Hyperbolic preferences, self-control problems, procrastination, retirement savings.

Timo Seppälä

Centre for Health Economics at STAKES
P.O.Box 22 (Lintulahdenkuja 4)
FI-00531 Helsinki
FINLAND
e-mail: timo.seppala@stakes.fi

* The author thanks *the Research Unit of Economic Structures and Growth (RUESG)* at the University of Helsinki and *the Yrjö Jahnsson Foundation* for financial support. I'm grateful to my advisors Erkki Koskela and Rune Stenbacka for helpful comments. I have also benefited from the insights of Suzanne Scotchmer, and participants at HECER labour and public economics seminar.

1 Introduction

Behavioural approach has offered new tools for consumption-saving economics to solve and explain some of the puzzles faced with conventional models. By using behavioural models some interesting and plausible solutions can be provided e.g. to undersaving, and overuse of credit cards. In this paper, we continue on this path of consumption-saving analysis by exploring existence impacts of a retirement savings program on different kinds of hyperbolic agents.

Börsch-Supan and Brugiavini (2001) list four developments why savings in general will be influenced forcefully in Europe: population ageing, increasing mobility, changing society's view about the role of the state in providing social security, and finally new financial landscape in Europe. They also state among others that one of the most important, if not the most important, interaction between savings and policy is centered on pensions. In a vast part of Europe large masses of employees are in the phase of transition from an employee to a pensioner. So far, quite many European countries have had pension funding as a part of their social security system, called usually as pay-as-you-go pension system. The system works in general so that people who are currently in the workforce provide pension for those who are retired. Recent analyses have shown that the mass which is to be in the retiring phase in few years is intolerable to the current, social security based, system.[1] It has been clearly stated in the public debate that a pension reform is required to accommodate new challenges in the financial environment of pension funding. There are basically three different suggested ways to route the current pension system. Firstly, complete dissolution of the public pension system by replacing it with a privately funded pension system. Secondly, partial dissolution of the public pension system while increasing workers' responsibility of pension funding by an amount it is reduced by the dissolution. Thirdly, saving in advance to cover the increasing future costs of the public pension system, i.e. increasing current payroll taxation. Intriguing question springs then up: what are the effects of the chosen pension funding method on individual saving behaviour.

This paper concentrates on the first of the given solution concepts. According to that pension funding would be arranged so that firms govern their workers' pension savings as it is currently the case in the United States.[2] Similarly, private funding could, of course, be just a part of individual

[1] For a very thorough analysis about current state in EU see e.g. Holzmann et al. (eds.) (2003).

[2] Retirement saving programs (or plans) are popularly known by the sections of tax code that address different kinds of employers: colleges and non-profits generally offer 403(b) plans, while private companies usually offer 401(k)s, and

saving as the administratively lightest option would suggest. In this paper we do not distinguish between instances that provide a savings program but we rather contrast a lucrative savings program with regular, checking account, saving. Lucrativety is provided by a higher interest rate in savings program than paid on checking account saving. The savings program and checking account are considered only possible options for an individual to save for his own pension in our set up where the public pension system is terminated. Our study compares individual consumption-saving behavior in two cases: (1) in the absence and (2) in the presence of a retirement savings program. We concentrate on possible consequences of hyperbolic behaviour on retirement savings when the public pension system is replaced completely with private saving. Our particular interest is on negative existence effects of a retirement savings program. We take ex-post-transition perspective, and hence we do not speak out how the overlapping phase between the old and the new pension system should be arranged.

We depart from the conventional consumption-saving analysis by assuming hyperbolic intertemporal preferences instead of (conventional) exponential preferences. This causes that the agent's action plans can be time inconsistent, i.e. the agent suffers from a *self-control problem* (SCP) à la O'Donoghue and Rabin (1999c).[3] Our main argument for using hyperbolic preferences is that a large body of evidence has shown that those fit the best in the data when completion of a task is studied. Several empirical studies have shown that people are eager to preproperate pleasant tasks and keen to procrastinate onerous tasks. Completion of a task, e.g. revising an article or enrolling on a retirement savings program requires effort that is costly in terms of utility but not usually in money. So far, the conventional models have been quite much incapable of sorting out the reasons for procrastination behaviour while hyperbolic models do well in explaining behaviour that is hard to explain plausibly with the conventional models.[4] In a pure consumption-saving set up, number of recent behavioral economics studies have shown that a model with hyperbolic preferences fits

governmental groups provide 457(b)s; there are other plans called 401(a)s and qualified plans that work in generally the same way. Note also IRAs, that are, from their part, tax-advantaged Individual Retirement Accounts.

[3]In general, hyperbolic preferences are such that the marginal rate of substitution between consumption at two future dates depends on the date at which evaluated. Strotz (1955-56) established that planned actions in an discrete intertemporal choice problem are time-inconsistent *if and only if* discounting is non-exponential *and* there is no commitment device available. Phelps and Pollak (1968), and Pollak (1968) continued developing Strotz (1955-56) work and build up the framework for non-exponential discounting and self-control problems. Then, Laibson (1994, 1996, 1997), Laibson et al. (1998) and Harris and Laibson (2001) formulated a solid theory field to study and apply the assumption of hyperbolic preferences in a consumption-savings analysis with multiselves approach. Micro-oriented studies from similar starting points were then presented in O'Donoghue and Rabin (1999a, b, c, 2001a, b, forthcoming).

[4]See e.g. Ainslie (1992), Loewenstein and Prelec (1992), and Frederic et al. (2002).

better also to consumption-saving data than exponential preferences do.[5] For example, a detected drop in consumption around the retirement was hard to explain with traditional life cycle models while use of a hyperbolic consumption-saving model gave a relief and solved this problematic issue with a plausible explanation. Hence, we lean on empirical and experimental data showing and reinforcing the matter that hyperbolic preferences do describe better an individual intertemporal decision process than what exponential preferences appear to do.

Our consumption-saving modeling follows strongly the lines of Laibson (1996) whereas the fashion of procrastination and agent's types are captured less intensively from O'Donoghue and Rabin (1999c, 2001b, forthcoming). However, our model is a stripped-down version of their models in some dimensions. Our model is simplified to gain a good tractability and a strong emphasis on the most important results. The time frame, on the other hand, is *finite* instead of infinite. With this choice we try to adopt a bit more realistic picture of individual retirement saving behaviour. This causes many concepts to be non-stationary, and hence structures and conditions become more complex than e.g. O'Donoghue and Rabin (2001b) have in their *infinite* time-frame model. The most easily seen difference deals with threshold values that O'Donoghue and Rabin (2001b) get for the agent's decision rules. For an infinite time frame, these thresholds do not vary in the passage of time. Contrary, in our setting such constant threshold values are absent due to the existence of terminal period. We simplify the model also by choosing not to have uncertainty. Further, the agent's only tasks are to make an intertemporal consumption-saving plan in every period, and if possible, also to choose whether or not and when (if ever) to start a lucrative retirement savings program. As assumed for most of savings program models, also in our model, enrolling on the savings program is assumed to be costly in terms of utility.[6]

Procrastination is possible when the savings program is present. Our procrastination analysis relates to O'Donoghue and Rabin (1999c) where a trasfer of a fixed size capital stock from one account to a better paid account is studied. However, in their model consumption behaviour is not contemplated at all as the agent does not have any possibilities to consume the capital stock or other wealth. The sole task for the agent is to transfer the capital stock from one to another account. We, instead, consider here a slightly different case where the agent has to make simultaneous decisions

[5] See e.g. Laibson et al. (1998)

[6] For reasoning and justification see e.g. Madrian and Shea (2001) and the preliminary work of Choi et al. (2004) that provide empirical studies about 401(k) retirement savings program and enrollment on it.

about optimal consumption and when to complete the task of starting a lucrative retirement savings program. Naturally, these two tasks are connected resulting in certain effects on each other. Our contribution springs up from a time-varying budget constraint causing non-stationary motives to start, and procrastinate starting of, the savings program. For the motives to procrastinate, flexibility in current budget constraint is important as well.

The results establish that depending on the agent's type, that are naive, sophisticated and learning naive, consumption-saving plans are consistent with realising actions only for the sophisticated agent while inconsistent for the two other types. The seminal result from Pollak (1968) stating that in the absence of a commitment device the realising actions for all the types of the agent, namely naive and sophisticated, with a logarithmic instantaneous utility function are the same, holds also in our model, and it extends to hold for the learning naive agent too. More fascinating and novel results are provided when the retirement savings program, SP from herein, is present. We find that if starting of SP is *worthwhile*, i.e. it is utility increasing even after suffering a non-monetary utility cost to enroll on it, the naive agent's retirement savings can be greater in the absence of SP than in the presence of it. The sophisticated agent gains while the learning naive agent can gain or suffer from the existence of SP in the terms of retirement savings. Whether the learning naive agent gains or suffers from the existence depends on the level of the effort cost. These results give us a possibility to discuss about optimal policy that should be exercised, and also about caveats of certain policies, if a planned pension reform would be a complete substitution of the publicly funded pension with private saving. Relating our results with the results from current literature gives then policy suggestions which establish that if enrollment in SP incurs utility cost, this utility cost should be decreased close to zero with non-automatic enrollment so that procrastination is not tempting. When lowering the effort cost is expensive relative to the cost of increasing individual's capability to learn about self-control problems, the public authority should invest in and emphasise the learning possibilities. Otherwise, an automatic enrollment on SP with optimised default should be used. Further more and intuitively, since very *existence* of SP can make the naive and learning naive worse off in terms of retirement time savings, it is also possible that the publicly financed system with some minor changes results in, after all, the best outcome. That is, if the publicly financed retirement system provided originally more than the savings program yields for the sophisticated agents, the publicly financed system could be maintained by decreasing the replacement rate on the respective of the sophisticated agents ability to save along SP and all the types should accepted it.

The rest of the paper is constructed as follows. In Section 2 we present the basic set up, while in Section 3 we analyse consumption and saving behaviour in the absence of SP. In Section 4 we incorporate the savings program into the analysis, and in Section 5 we solve numerically some results to give support for formal analysis. Finally, Section 6 concludes.

2 Basic Set Up

This section introduces first our basic set up in which modeling of consumption-saving follows strongly the lines of Laibson (1996). We also give a formal definition for different agent types in terms of agent's awareness about his *future* SCPs. The verbal counterpart of our formal type classification relates to O'Donoghue and Rabin (1999c, 2001b, forthcoming). The basic set up as a whole is thus a modification of the models established in Laibson (1996) and O'Donoghue and Rabin (1999c, 2001b, forthcoming). After releasing the basic set up with necessary tools we exploit it to contemplate consumption-saving behaviour of the different agent types. Finally, consumption-saving behaviour is compared between the different types and across separate selves of the agent.

From now on we suppress notation by denoting any subsequence $\{n, n+1, ..., K\}$ of natural numbers, $\mathbb{N}$, by $\mathbb{N}_n^K$, and whenever $n = 1$, i.e. for $\mathbb{N}_1^K = \{1, 2, ..., K\}$, we use $\mathbb{N}_1^K \equiv \mathbb{N}^K$. For other notational ease we omit, whenever it is possible, those super and subscripts that are obvious and clear from the context, or redundant for the analysis.

Time is discrete and finite, it runs from period $t = 1$ to period T, hence $t \in \mathbb{N}^T$ for all t. We use the multiselves approach, and thus we model the agent as a sequence of autonomous temporal selves.[7] These selves are indexed by the respective periods in which they control the choice variable, hence the set of selves is $\mathbb{N}^T$ and we denote a self at period t by self(t). For all i and for all $t \in \mathbb{N}^{T-1}$, agent's intertemporal preferences are assumed to be hyperbolic and we present them

[7] We say that discounting is quasi-hyperbolic since discounting is exponential for all periods from $t+1$ onwards and captures only the key qualitative property of hyperbolic discount function, i.e. a faster rate of decline in the short run than in the long run. From now on we always refer to *quasi-hyperbolic* discounting with the term *hyperbolic discounting.*

In the multiselves model agent's identity is cut in as many selves as there are periods in the model, i.e. periods that the agent lives. Then, the *present self* plays an extensive form game with his *future selves* with a strategy that depends on agent's type.

with the following modification of the model proposed in Phelps and Pollak (1968).

$$U_t^i \equiv u(c_t) + \beta_{(t,i)} \sum_{\tau=1}^{T-t} \delta^{\tau} u(c_{t+\tau}), \tag{1}$$

where U_t is the agent's time-additive intertemporal utility at time $t \in \mathbb{N}^{T-1}$ that consists of the time-separable, continuous and strictly concave instantaneous utility functions, $u(c_t)$, and consumptions, c_t, and naturally, $U_T = u(c_T)$. The discount parameters are $0 < \beta < 1$ and $0 < \delta < 1$, and have the following interpretations: β is the short term, sometimes called as a time-inconsistent, discount factor whereas δ is conventional long term exponential discount factor. The instantaneous utility function $u(c_t)$ is assumed to be in the class of constant relative risk aversion ($CRRA$). Its general representation is thus $u(c_t) = \frac{(c_t)^{1-\rho}-1}{1-\rho}$, for which $\rho > 0$ ($\rho < 0$) implies that the agent is risk averse (risk lover). In the analysis we fix $\rho = 1$, and hence we use the explicit form of $u(c_t) = \ln c_t$.[8] In Def.(1) we have reserved the possibility of having different βs for different selves and for the different agent types. That is $\beta_{(t,i)}$ denotes β for self(t) of type i. We then say that the smaller (the bigger) the time-inconsistent discount factor β is the bigger (the smaller) agent's SCP is.

We have three different agent types: the naive agent (NA), the learning naive agent (LN), and the sophisticate agent (SA), hence $\forall i$ holds $i \in \mathcal{I} \equiv \{NA, LN, SA\}$. The agent is naive when he anticipates that his preferences do not change in time, i.e. he is not *aware* neither does he *know* about his future preferences reversals.[9] The agent is sophisticated when he *knows* his future preferences. The agent is learning naive when, after the first period, the agent is *aware* about his self-control problem but underestimates the magnitude while this failure gets smaller as he learns his preferences along the passage of time. Formally, we define.[10]

Definition 1 *Let $\beta_{(t+j)}^{(t)}$ denote self(t)'s perceptions about β_{t+j}, and let the true state be such that $\beta_{t+j} = \beta_t = \beta < 1 \,\forall t \in \mathbb{N}^T$ and $\forall\ j \in \mathbb{N}^{T-t}$. Then*
a) the agent is naive iff $\beta_{(t+j)}^{(t)} = 1\ \forall t \in \mathbb{N}^{T-2}, \forall j \in \mathbb{N}^{T-t}$;
b) the agent is learning naive iff $\beta_{(1+j)}^{(1)} = 1\ \forall j \in \mathbb{N}^{T-1}$ and $\beta < \beta_{(t+j+1)}^{(t+1)} < \beta_{(t+j)}^{(t)} < 1\ \forall t \in$

[8] It is easy to verify for $u(c_t) = \frac{(c_t)^{1-\rho}-1}{1-\rho}$ that it is continuous in almost everywhere, i.e. it is continuous everywhere except at $\rho = 1$, for which we get as a special case $u(c_t) = \ln c_t$. Then $u'(c_t) > 0$, $u''(c_t) < 0$ for all $c_t > 0$ and for all $\rho > 0$, and hence it is strictly concave and it fulfills the property of diminishing marginal utility. Due to the absence of uncertainty and non-trivial borrowing constraints in our model, a non-linear utility function enables us to have also other than corner solutions.

[9] For a generic individual we use noun "*he*". The choice was made by tossing a fair coin.

[10] The reasoning and deduction of the definition is consigned to Appendix A.

$\mathbb{N}_2^{T-2}, \forall j \in \mathbb{N}^{T-t}$;

c) the agent is sophisticated iff $\beta_{(t+j)}^{(t)} = \beta \ \forall t \in \mathbb{N}^{T-2}, \forall j \in \mathbb{N}^{T-t}$.

We assume $0 < \beta_{(t,i)} < 1$ and $\beta = \beta_{(t,i)}$ for all $(t,i) \in \mathbb{N}^{T-1} \times \mathcal{I}$, and with out loss of generality we set $\beta_{(T,i)} = 0$ for all i since the terminal self does not have relevant intertemporal preferences.

The agent starts working in period 1 with a certain and constant periodical disposable salary of w_t.[11] The agent knows that he will retire in period $P+1$ and that he is forced to save for his own pension as there are no incomes during the pension. This is to say that $w_t = w > 0$ if $t \in \mathbb{N}^P$ and $w_t = 0$ if $t \in \mathbb{N}_{P+1}^T$.

The agent's sole every period problem is to maximise his intertemporal utility U_t^i by choosing consumption (and hence saving) subject to his budget constraint. To complete the task, the agent makes a consumption-saving plan over the remaining periods that maximises his intertemporal utility. To keep the analysis simple we do not set any *non-trivial borrowing constraints* for the agent.[12] This implies that the agent's liquidity constraint is always the sum of present value of the agent's labour income and current financial wealth. Hence, the agent makes his consumption-saving choices subject to the budget constraint that can be defined recursively as follows.

$$c_t \leq W_t \equiv (W_{t-1} - c_{t-1})R,$$

where net wealth W_t includes the discounted present value of the stream of labor income added to the sum of current financial wealth, i.e.

$$W_t = Rs_t + w \sum_{\tau=0}^{P-t} \delta^\tau,$$

and where R is the constant gross return on wealth with the interest rate r, i.e. $R = 1+r$. Notably, the interest rate on both, saving and borrowing, is r. About the long-term discount factor δ we assume for simplicity that it is the inverse of the gross interest rate $\delta = R^{-1}$. The agent does not have initial savings and we assume that he does not leave any bequest, formally this means $W_1 = w \sum_{\tau=0}^{P-1} \delta^\tau$ and $W_T - c_T = 0$.

[11] We call w_t as *disposable salary*, since it is considered as net-of-tax salary.

[12] With a non-trivial borrowing constraint we refer to the constraint that would be lower than the agent's current wealth.

To derive the optimal consumption-saving plan, self(t) of type i has to solve the following intertemporal utility maximising problem.

$$\max_{\{c_j\}_{j=t}^{T-t}} U_t^i$$

$$s.t. \qquad \sum_{\tau=0}^{T-t} \delta^\tau c_{\tau+t} \leq W_t.$$

The set up is finite in time, and the agent is assumed to be able to conduct backwards induction for the maximisation problem. Consequently, the solution will be a subgame perfect equilibrium from the viewpoint of the self who does the maximisation in his turn.[13]

Finally and before turning to analyse consumption-saving behaviour, we still present some useful concepts and notation. We let $C_i^{t*} \equiv \{c_{t+j}^{t*,i}\}_{j=0}^{T-t}$ denote a sequence of *planned* consumption choices of self(t) of the type i, whereas $C_i^t \equiv \{c_{t+j}^i\}_{j=0}^{T-t}$ denotes a sequence of *realised* consumption choices from the period t onwards. We denote $C_i^{t*} = C_i^t$ if and only if $c_{t+j}^{t*,i} = c_{t+j}^i$ for all $j \in \mathbb{N}^{T-t}$ otherwise $C_i^{t*} \neq C_i^t$.

Saving resolves then as a residual to consumption, and hence what is not consumed is saved. Let s_{t+1}^i denote a realised saving from period t to period $t+1$ in terms of the valuation of period t. Thus, we let $S_i^{t*} \equiv \{s_{t+j}^{t*,i}\}_{j=0}^{T-t}$ denote a sequence of self(t)'s *planned* periodical saving and respectively $S_i^t \equiv \{s_{t+j}^i\}_{j=0}^{T-t}$ denotes a sequence of *realised* saving. Again, we denote $S_i^{t*} = S_i^t$ if and only if $s_{t+j}^{t*,i} = s_{t+j}^i$ for all $j \in \mathbb{N}^{T-t}$ and $S_i^{t*} \neq S_i^t$ otherwise.Realised temporal saving during the working periods is clearly $w - c_t + Rs_t \equiv s_{t+1}$ and during the retirement time it is $Rs_t - c_t = W_t - c_t \equiv s_{t+1}$. From the assumption of '*only trivial borrowing constraints*' follows that s_t may be negative or positive during the working periods. Particularly, self(t) is a debtor whenever $s_t < 0$ or $s_{t+1} < 0$ or both holds. Clearly, self(t) borrows always if his consumption exceeds his current savings, i.e. if $c_t - w > 0$ and $c_t - w > Rs_t$. There are then two possible cases when self(t) borrows: 1) self(t) consumes more than earns and has insufficient non-negative savings, i.e. $c_t - w > Rs_t \geq 0$, 2) self(t)

[13] Whether or not the solution will be subgame perfect equilibrium also from other selves viewpoint depends completely on each self's correctness of perceptions about other selves behaviour. Naturally, when these perceptions are incorrect it might happen that even though some self(t) considers his choice as a part of subgame perfect equilibrium, due to fallacious perceptions future selves do not behave as he thought and he might have chosen differently if he had known the true behaviour.

consumes more than earns and his current savings are non-positive, i.e. $c_t - w > 0 \geq Rs_t$.[14],[15]

3 Consumption and saving behaviour

3.1 Consumption

Next proposition reveals us both the realising consumption choices and the consumption plan of each type and self.

Proposition 1 *Given the basic set up,*
(i) for any $i \in \mathcal{I}$, and any $t \in \mathbb{N}^{T-1}$, the realised consumption path is

$$C_i^t = \{\iota_j W_j\}_{j=t}^T,$$

where $W_j = Rs_j + w_j \sum_{\tau=0}^{P-j} \delta^\tau$, and $\iota_j = \left(1 + \beta \sum_{\tau=1}^{T-j} \delta^\tau\right)^{-1}$ $j \in \mathbb{N}^T$.
ii) the relation between a realising consumption path C_i^t and a plan C_i^{t*} of type i of self(t) is

$$\begin{aligned} C_{NA}^{t*} &= \{\iota_t W_t, \underbrace{\beta\iota_t W_t, \beta\iota_t W_t, ..., \beta\iota_t W_t}_{T-t\ elements}\} \neq C_{NA}^t \text{ for any } t \in \mathbb{N}^{T-2}, \\ C_{LN}^{t*} &= \left\{\iota_t W_t, \left(\iota_{(j)}^{(t)} W_j^{t*}\right)_{j=t+1}^T\right\} \neq C_{LN}^t, \\ C_{SA}^{t*} &= \{\iota_j W_j\}_{j=t}^T = C_{SA}^t. \end{aligned}$$

Proof. All proofs are consigned to Appendix B. ■

The claim (i) in Proposition (1) gives the realised consumption path. We find that the agent consumes in every period a certain fraction of his wealth and this fraction is self's time-dependent marginal propensity to consume which is linear in current wealth. Moreover, the realising consumption path is independent of agent's type, which is a consequence to our choice of the logarithmic instantaneous utility function.[16] The claim (i) in Proposition (1) has some useful properties. Firstly,

[14] Later on, we will show that in the basic set up and for the all agent types only the second possibility occurs but never the first one.

[15] Note that we count the repayment of a loan in the category of positive saving, and hence we consider *self(t)* is saving also in the case where the savings are still negative but increasing, i.e. $s_{t+1} \leq 0$ and $w - c_t > 0$.

[16] This result was first established and proved in Pollak (1968). In the mentioned paper a theorem states that "...in the N decision point case, if the instantaneous utility function is logarithmic, then the naive optimum path and the sophisticated optimum path coincides."

since the realising consumption choice is independent of the agent type, we immediately know that also the realising saving decisions must be independent of the type. Secondly, consumption-saving plan comparisons across the agent types become comfortable and easy, since we know that wealth, consumption and saving in a planning period are the same regardless of the agent's type. From the claim (ii) we see that planned consumption coincides with realising consumption only for *SA* while *NA* and *LN* systematically makes an error with their consumption plans. We find that every self of *NA* consumes more during the planning period than what he plans to consume during the subsequent periods, and that the planned consumption is constant in terms of the current valuation for the subsequent periods.[17] *LN* makes a similar mistake with *NA* when planning but his planning error decreases in time due to ability to learn from his past behaviour.[18] Only *SA*'s every self(t) knows that future selves are not going to honor his current preferences, and thus makes his consumption plan conditionally on correctly anticipated future behaviour.

Before starting with the saving analysis we present one lemma that will alleviate it. Until now, we know that the sequence of realised consumption of any type of the agent is of the form $C^t = \{\iota_t W_t\}_{j=t}^{T}$. However, we do not know whether the path of consumption is monotonic or not. Even though we know that $W_t > W_{t+1}$ $\forall t \in \mathbb{N}^{T-1}$, and hence, wealth decreases monotonically in time, $sign\,\{c_t - c_{t+1}\}$ remains unknown, since at the same time $\iota_t = \left(1 + \beta \sum_{\tau=1}^{T-t} \delta^{\tau}\right)^{-1}$ increases monotonically in time, i.e. $\iota_t < \iota_{t+1}$ $\forall t \in \mathbb{N}^{T-1}$. So, in principle it could happen that wealth decreased proportionally less than ι_t increased resulting in increasing consumption path, or vice versa, or both at different times. To this end, the following lemma solves this problem by stating that realised consumption decreases monotonically in time, i.e. $c_t > c_{t+1}$ $\forall t \in \mathbb{N}^{T-1}$.

Lemma 1 *Given the basic set up,* $c_t > c_{t+1}$ $\forall t \in \mathbb{N}^{T-1}$.

[17] *Self(t)*'s consumption plan for periods $t + j \in \mathbb{N}^{T-t}$ has straight forward analogy to exponentially discounting agent's plan. It is commonly known that when condition $\delta = R^{-1}$ holds with intertemporal preferences, where instantaneous utility function is logarithmic the optimal consumption is given by $c_t^* = c_{t+j}^*$ $\forall j \in \mathbb{N}^{T-t}$. Now, since *NA* considers that he will behave as an exponential discounter in the future periods his plan for those periods is given by the same rule that an exponential discounter would use.

[18] Of course, self(t) of any type would prefer the case where the later selves honored self(t)'s preferences to the case where every self($t + j$) maximises his own intertemporal preferences. To attain the consumption choices that would follow self(t)'s intertemporal preferences is now, however, impossible because of the absence of commitment devices.

3.2 Saving

Recalling that we defined saving as a residual of consumption enables us to derive saving behaviour for the different agent types by using our results from the previous analysis, namely using Proposition (1). We are primarily interested in agent's saving for the retirement, i.e. $s_{P+1} = R^{-1}W\left(P+1\right)$. Among lifetime saving behaviour the next proposition specifies also differences and similarities of planned and realised retirement savings between the different agent types.

Proposition 2 *Given the basic set up,*

i) $S_i^t = S_j^t$ *for all* $i, j \in \mathcal{I}$, *and for all* $t \in \mathbb{N}^T$,

ii) $S_i^{t*} \neq S_j^{t*}$ *for all* $i, j \in \mathcal{I}$ *such that* $i \neq j$, *and for all* $t \neq 1$, $t < T-1$,

iii) $s_{P+1}^{t*,i} \neq s_{P+1}^{t*,j} \geq s_{P+1}$ *for all* $i, j \in \mathcal{I}$ *such that* $i \neq j$, *and for all* $t \neq 1$, $t < P$,

iv) $\Delta s_{P+1}^{t*,LN} \leq \Delta s_{P+1}^{t*,NA}$, *for all* $t \leq P$ *where* $\Delta s_{P+1}^{t*,i} = s_{P+1}^{t*,i} - s_{P+1}$ *for* $i \in \{NA, LN\}$.

The claim (i) in Proposition (2) follows directly from the claim (i) in Proposition (1). Since the agent's realising consumption is independent of the agent type and it is the same across the types, it must be the case that also the sequences of realising saving choices are identical across the different agent types. According to the claim (ii) the saving plans are all different between the different agent types. Particularly, the claim (iii) states that the agent's plans about total retirement savings are different between the different agent types. Furthermore, by using the analysis behind the claims (i) and (ii) with the claim (iii) we get to know that the plans about total retirement savings fail to hold for *NA* and *LN* whereas *SA* knows perfectly his funds for the pension. Contrary to *SA*'s perfect knowledge, *LN*'s and *NA*'s self($t < P$) projects the plans about retirement savings higher than what those ever will be. *LN*, however, learns that his preferences are subject to change in the future periods, and hence his plans about the retirement savings approach *SA*'s plan. *NA*'s plans approach to the actual amount of retirement savings also but convergence is solely caused by decreasing wealth.

Let us now consider agent's borrowing behaviour. Like it was stated earlier borrowing can emerge, in principle, in the two different cases where in both consumption must exceed the wage: 1) the agent does not have any current positive savings, i.e. $Rs_t \leq 0 < c_t - w$; 2) the agent has insufficient positive savings to cover the wage exceeding consumption, i.e. $0 < Rs_t < c_t - w$. The next proposition facilitates our the analysis by cancelling out the latter state of the emerging borrowing possibilities. In other words, Proposition (3) establishes that given the basic set up the agent never consumes more than his current *positive* savings is. The proposition is presented in more general form, and it actually states that if the agent starts to accumulate funds during the working periods, he accumulates funds from that period onwards until the retirement, and hence does not borrow ever again. The proposition states also that the agent borrows money in every period before the first period of positive saving.

Proposition 3 *Given the basic set up,*
i) If $s_{t+1} > Rs_t$ *for some* $t < P$ *then* $w - c_{t+j} > 0 \ \forall j \in \mathbb{N}^{P-t}$,
ii) If $t = \min\{k \leq P \,|s_{k+1} \geq Rs_k\}$ *then* $s_{t-j} \leq 0$ *and* $w - c_{t-j} < 0 \ \forall j \in \mathbb{N}^{t-1}$.

By Proposition (3) we know now that it never happens that the agent first accumulated funds and then would consume more than his positive savings is, i.e. it is not possible that $Rs_t > 0$ and $w - c_t + Rs_t < 0$.

In the next section we turn to a context where the agent is able to start a fixed-contribution retirement savings program, and hence in principle a weak commitment device is introduced into the analysis. We will see that then the sophisticates and the learning naives start to gain from their knowledge.

4 Saving Program and Procrastination

We now introduce a retirement savings program into the analysis. To get started, let us assume the basic set up, and let us expand it by assuming that instead of having possibility to *regular saving*, i.e. saving in the fashion of previous sections, the agent has also a possibility to make a costly effort to find a supplier of a fixed-contribution retirement savings program (SP) that guarantees him the retirement savings of B_l if the program is started in period l.[19] In other words, we assume that the agent knows, $\forall i \in \mathcal{I}$, that by making the costly effort he can ensure retirement savings of size B_l by saving a fixed-contribution, γ, along SP started in period l. Notation B_l^t then denotes (i) SP that is available to self(t), and (ii) its size in the terms of money in the first pension period, $P + 1$, if the program were started in period l.

Within this framework our interest is then to find out how long the agent of the type i possibly procrastinates until he makes the effort to start the retirement savings program. Our other interest

[19] A costly effort could just simply mean, for example, going to a bank and enroll on SP, where the cost is the harm the agent feels when he has to quit watching a football match from the televison.

is to find out how planned and realised retirement savings then differ across selves. Finally, we are interested in contemplating differences in agent's behaviour in the presence and absence of SP.

4.1 Revising the basic set up

Effort cost, utility, and budget constraint We assume that making the effort is costly, and that the effort cost is an immediate non-monetary one-shot cost. Its effect is assumed to be negative on the total utility and it has the size of $u(e)$ in terms of the instantaneous utility. The effort cost e is, thus, a money-metric measure that enables us to measure, in terms of the utility, a harm that the agent feels when completing the task of starting SP.[20] We also assume that the effect of the effort cost is additively separable in the agent's intertemporal utility function. Hence, if the agent made the effort, let us say, in period $t = 2$, the utility that self(2) would get in that particular period would be then the instantaneous utility from the consumption subtracted by the disutility from making the effort, i.e.

$$u(c_2) - u(e),$$

and the corresponding intertemporal utility from the viewpoint of self(1) would be then given as

$$U_1 = u(c_1) + \beta\delta\left(u(c_2) - u(e)\right) + \beta\delta^2 u(c_3) + \cdots + \beta\delta^{T-1}u(c_T).$$

Let l denote the period when the agent makes the effort. Since l is unknown we modify the definition of the intertemporal utility function given in Eq.(1) as follows.

$$U_t = u(c_t) - k_t u(e) + \beta\sum_{\tau=1}^{T-t}\delta^\tau\left(u(c_{\tau+t}) - k_{\tau+t}u(e)\right),$$

where

$$k_t = \begin{cases} 0, & if\ t \neq l \\ 1, & if\ t = l \end{cases}.$$

Then, by making the effort in period l self*(l)* commits himself and tries to commit all his subsequent selves to contribute the fixed SP contribution γ in every period from period l onwards until to the

[20] It is important to remark that due to our definition of the effort cost, it will not show up in the budget constraint but only in the intertemporal utility function.

terminal working period P. If all selves from l to P contributes on SP, the agent's intertemporal budget constraint for self(t) can be written as follows.

$$c_t + \sum_{\tau=1}^{T-t} \delta^{\tau+1} c_{t+\tau} \leq Rs_t + \sum_{\tau=0}^{P-t} \delta^{\tau}(w - d_{\tau+t}\gamma) + \delta^{P+1-t} B_l \equiv W_t^{B_l}, \tag{2}$$

where

$$d_n = \begin{cases} 0, \ if \ n < l \leq P \\ 1, \ if \ n \geq l \end{cases}$$

and $W_t^{B_l}$ denotes a *considered wealth* in period t if SP were started in period l. Notice that we use the term 'considered wealth', since self(t) *considers* that he will start SP in period l, and thus if self(t) is confident about starting SP in the considered period he naturally includes this part of increased wealth in his budget constraint when forming a consumption-saving plan.

Feasibility and cover It is trivial that the agent would not be interested in SP if it caused a loss in his life time budget. Hence, for any SP to be feasible it is necessary that starting the program will not cause a loss in the agent's total budget. Naturally, feasibility is guaranteed for any savings program for which $R_B = (1 + r_B) > R = (1 + r)$, where r_B is the interest rate paid along the program and r is the interest paid on the regular savings. From now on, we assume that the available savings program is of the form

$$B_l^t \equiv \gamma \sum_{\tau=l-t}^{P-t} R_B^{P-\tau-t+1}, \tag{3}$$

where $R_B = 1 + r_B > 1 + r = R$, and hence $r_B > r$. Setting $B_l^t = \gamma \sum_{\tau=l-t}^{P-t} R_B^{P-\tau-t+1}$ cancels out other bonuses that a bank could pay for the agent who starts SP. The agent only receives better than the regular interest rate paid on the contributions γ in SP.[21]

We say that SP has *l-cover* if SP that is started in period l guarantees the retirement savings greater than or equal to the retirement savings self*(l)* of the type i anticipates being able to save

[21] We remark here that any feasible saving program is such that when it is started in period l, later selves will not have an incentive to exit the program. This property and its reasoning will be discussed more carefully later on.

without the savings program. Hence, SP has t-cover whenever the Eq.(4) below is satisfied.

$$B_t^t \equiv \gamma \sum_{\tau=0}^{P-t} R_B^{P-\tau-t+1} \geq R s_{P+1}^{t*,i}. \tag{4}$$

By elaborating Eq.(4) we get the lower bound for the contribution γ that satisfies the property of t-cover for any given and fixed $r_B > r$, i.e.

$$\gamma \geq \frac{R s_{P+1}^{t*,i}}{\sum_{\tau=0}^{P-t} R_B^{P-\tau-t+1}}.$$

In other words, the property of t-cover fixes the lower bound of the contribution to a certain level for any given interest rates r and r_B for which $r_B > r$.[22] We will from now on assume that the available savings program is of the form *'1-cover for NA'*, and hence for given r and r_B for which $r_B > r$ we set γ as follows[23]

$$\gamma \equiv \frac{R s_{P+1}^{1*,NA}}{\sum_{\tau=0}^{P-1} R_B^{P-\tau}}. \tag{5}$$

We make the assumption '*1*-cover for *NA*', since by that assumption we fix γ, r_B, and r at the *1*-cover level for all $i \in \mathcal{I}$. From the preceding saving analysis we know that $s_{P+1}^{1*,NA} \geq s_{P+1}^{1*,i}$ for all $i \in \{LN, SA\}$. So, for the fixed r and r_B we know that if SP has 1-cover for NA it certainly has 1-cover for LN and SA. Due to fixing γ, r_B, and r we may from now on drop the superscript from the notation of SP, and hence B_k^t becomes B_k, where B_k denotes SP that is started in period k and the rate of interest paid on the contribution as well as the size of contribution are fixed to the level that satisfies *1*-cover for NA.

Worthwhile savings program, and procrastination Any savings program with any fixed positive γ and any $r_B > r$ would be certainly always worthwhile if agent's possibilities to borrow against the net present value of wealth were not constrained *and if* making the effort were *not* costly, i.e. if $u(e) = 0$. However, from the assumption $-u(e) < 0$ it follows that SP has to be *worthwhile* in a such sense that the utility loss from making the effort will be covered with increased consumption

[22] Of course, one could as well fix γ first, and after that derive such interest rates r_B and r that defined properties of an l-cover would be satisfied.

[23] By setting the contribution γ as given in Eq. (5), we get for our numerical calibration the contibution to be approximately 3% of the salary. This is a common default in 401(k) retirement saving programs. See e.g. Madrian and Shea (2001), and Choi et al (2004a).

possibilities due to SP.

To illustrate this property let us consider two different cases, one where the agent makes and the other where he does not make the effort. Let $C_i^{t*,B_t} \equiv \{c_j^{t*,B_t}\}_{j=t}^T$ denote self(t)'s consumption plan when SP is available, and assume that self(t) makes the effort in period t. Clearly, $c_j^{t*,B_t} > c_j^{t*}$ for all $j \in \mathbb{N}_t^T$, for any $\gamma > 0$, and for any $r_B > r$ regardless of the size of $u(e)$. The planned consumption increases since self(t)'s discounted lifetime income has increased by the amount of $\gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t}\delta^\tau\right)$, and so, self($t$)'s lifetime budget constraint has moved "outwards". Since we have given to the agent a possibility to smooth out his consumption over the life time, he will also do so. The incremental income will be then dispersed over the life time left according to the agent's intertemporal preferences. Thus, there exists incremental consumption ρ_j such that $c_j^{t*,B_t} = c_j^{t*} + \rho_j$ for $j \geq t$.[24] Then, denoting $U_{t*}(C_i^{B_t}) \equiv U_t(C_i^{t*,B_t})$ and respectively $U_t(C_i^*) \equiv U_t(C_i^{t*})$ neither *1*-cover nor the feasibility is sufficient to define the sign of the difference between the intertemporal utilities with and without the savings program, i.e.

$$sign\left[U_{t*}(C_i^{B_l}) - U_t(C_i^*)\right], \tag{6}$$

which is eventually crucial when the agent solves whether starting of SP is *worthwhile* at all ($sign[\cdot] > 0$ for some l) or not ($sign[\cdot] \leq 0$ for all l).[25] For any given SP the sign of the Eq.(6) is solely determined by the size of e as we will show next.

Now, $U_t\left(\{c_j^{t*} + \varepsilon_j\}_{j=t}^T\right) > U_t\left(\{c_j^{t*}\}_{j=t}^T\right)$ for any $\varepsilon_j > 0$. On the other hand, $\varepsilon_j > 0$ if and only if the agent considers making the effort in some period $l \geq t$.[26] Hence, in the case where self(t) makes the effort in the period t his intertemporal utility U_t according to his plan can be written as follows.

$$U_{t*}\left(C_i^{B_t}, u\left(e\right)\right) = U_t\left(\{c_j^{t*} + \rho_j\}_{j=t}^T, u\left(e\right)\right) = \hat{U}_t\left(\{c_j^{t*} + \rho_j\}_{j=t}^T\right) - u(e) = U_t\left(\{c_j^{t*}\}_{j=t}^T\right) \text{ for some } u(e),$$

[24] For example for NA in our setup planned incremental consumption is given by the rule $\rho_t = \beta^{-1}\rho_{t+i}$, and hence according to the agent's preferences the incemental consumption is the same for all periods from period $t+1$ onwards and $\iota_t \Delta W_t$ for the current period t, where $\Delta W_t = W_t^{B_t} - W_t$.

[25] We include the equality in the case of not considering SP worthwhile, and hence we make an implicit assumption that if there is no gain no loss from starting SP the agent prefers not to make the effort.

[26] Recall that $U_t(C_t^{B_l*})$ and $U_t(C_t^{t*})$ are just the imaginary intertemporal utilities that would be realised if the subsequent selves were able to stick to self(t)'s consumption-saving plan. Whether he really is capable of abiding by his consumption plan depends completely on his type as we saw in the last section.

where we have decomposed the intertemporal utility in two pieces by defining

$$\hat{U}_t(\cdot) \equiv U_t(\cdot, e) + u(e),$$

where $\hat{U}_t(\cdot)$ is the pure intertemporal utility from consumption.[27]

Clearly, SP gives the higher incremental incomes the sooner it will be started. On the other hand, from the viewpoint of self(t) , SP yields the smaller utility loss from making the effort the later program is started. Hence, the total effect from making the effort on the intertemporal utility depends on the relative rate of depreciation between a utility loss and a loss in incomes due to procrastination of starting SP. Then, SP is *worthwhile* from self(t)'s perspective if and only if

$$\min\left\{ k \geq t | \, sign\left[U_{t*}(C_i^{B_k}) - U_t(C_i^*) \right] > 0 \right\} \leq P.$$

In words, self(t) would like to start SP as soon as possible but dislikes the utility loss caused by making the effort, and hence he has a temptation to procrastinate it. If there is no such period before retiring that the utility gain from SP outweighs a utility loss from making the effort, self(t) abandons SP for good and all the results concerning consumption-saving behaviour are analogical from that self onwards with the analysis in the previous sections. For worthwhileness and attractivety we get then the following formal definitions.

Definition 2 *For self*$(t)\,, t \in \mathbb{N}^P$*, a savings program is*
i) worthwhile if for some $h \in \mathbb{N}_t^P$ *holds sign* $\left[U_{t*}(C_i^{B_h}) - U_t(C_i^*) \right] > 0$;
ii) h^{t*}*-worthwhile if* $h^{t*} \equiv \min\left\{ h \in \mathbb{N}_t^P \middle| \, sign\left[U_{t*}(C_i^{B_h}) - U_t(C_i^*) \right] > 0 \right\}$;
iii) attractive if $h^{t*} = t$.

Period h^{t*} is the first period when making the effort, and thus, SP is worthwhile from self(t)'s perspective.[28] Next, we will make a highly simplifying assumption about agent's behaviour, but we consider it plausible and reasonable due to agent's shortsightedness.

[27] We say 'pure' intertemporal utility, since the effect of making the effort is purged from $\hat{U}_t(\cdot)$ by adding $u(e)$ to the real intertemporal utility $U_t\,(\cdot, u(e))$.

[28] Our definition of the worthwhileness is analogous with the definitions of $\beta-$*worthwhile* and $\beta-$*best task* proposed in O'Donoghue and Rabin (2001b). However, due to slight differences in the set ups we use the term "worthwhile for self(t)" to describe the $\beta-$worthwhile task and the term "$h^{t*}-$worthwhile for self(t)" to describe the $\beta-$best task.

Assumption 1 *If $h^{t*} > t$, then $C_i^{t*,B_{h^{t*}}} = (C^{t*} \mid W_t)$; if $h^{t*} = t$, then $C^{t*,B_{h^{t*}}} = \left(C^{t*} \mid W_t^{B_{p^{t*}}}\right)$, where $(C \mid W)$ denotes the consumption plan C conditional on wealth W ,and p^{t*} denotes the optimal starting period for self(t) from his perspective.*

Assumption (1) says that if SP is unattractive for self(t), he considers that it will not be attractive for any subsequent self either, and hence he makes the consumption-saving plan conditional on his net present value of *real wealth*. Self(t) considers that since he would not start SP in the current period, there is no reason for him to assume that he will start it in the future. This kind of feature of behaviour can be interpreted as procrastinated *decision making*, i.e. the agent delays making his decision about making the decision.[29] From now on we refer to delayed decision making by *second order procrastination*. Assumption (1) says also that if, on the other hand, SP is attractive for self(t), he then makes the plan about the optimal starting period and in addition forms the consumption plan conditional on the present value of *considered wealth*.[30]

Take now the viewpoint of self(t) and consider some attractive savings program for him. It is important to notice that even though SP is attractive for self(t) it does not mean that the current period would be the optimal period to start SP from self(t)'s viewpoint. From self(t)'s perspective it would be optimal to start the attractive SP in period p that solves $\max_{\{p\}_{p=t}^{P}} U_{t*}(C_i^{B_p})$ conditional on that self(p) abide by self(t)'s starting plan. In general, for any self(t) and for any attractive savings program an optimal starting period will be chosen subject to perceptions about future behaviour. Without loss of generality and to ensure the uniqueness of the solution from self(t)'s viewpoint, we assume that if there are several ps that satisfy the maximisation, then self chooses the minimum from the set of possible options.[31] We can formally define the optimal starting period from the viewpoint of self(t) as follows.

Definition 3 *Self(t)'s optimal starting period, p^{t*}, for an attractive savings program satisfies*

$$p^{t*} \equiv \min\{\, p \in \mathbb{N}_t^P \,\big|\, \max U_{t*}(C_i^{B_p}) \text{ s.t. perceptions about future behaviour}\}.$$

[29] This kind of delaying of decision making is not peculiar. It is quite common for one to say that he needs a while before he is ready to make his final decision about the subject of concern.

[30] The second part of the assumption is discussed in Section 5. See also Hsieh (2003) for a closer analysis of anticipated income changes.

[31] This assumption implies that the agent is not tempted by unnecessary procrastination. If there are several same intertemporal utility resulting possibilities to start the attractive saving program self(t) considers that the sooner the better. Note also that this assumption is not contradictory to the assumption which defines the interface between worthwhile and not-worthwhile, and says that the agent considers the saving program not-worthwhile if it gives the same utility as never starting the saving program gives.

Then, adopting the fashion of procrastination from O'Donoghue and Rabin (2001b), we say that self(t) procrastinates whenever an attractive savings program exists but self(t) prefers starting of it in some later period to starting it in the current period.[32] The formal definition follows.

Definition 4 *Self(t) procrastinates if* $t < p^{t*}$.

Of course we do not say that the agent procrastinates if there is no worthwhile savings program available for any self(t), i.e. if $\{h|\, sign\left[U_{t*}(C_i^{B_h}) - U_t(C_i^*)\right] > 0\} = \varnothing$ for all $t \in \mathbb{N}^P$ and for all $h \in \mathbb{N}_t^P$. Hence, no procrastination by definitions emerges if there is no 1st order procrastination.[33]

4.2 Procrastination analysis

Given the revised set up, the intertemporal utility maximisation problem can be now written in the following form for any *non-retired* self(t) of any type $i \in \mathcal{I}$.[34]

$$\max_{\{c_t\}_{t=t}^T, l} U_t^i = \ln\left(\frac{c_t^i}{e^{k_t}}\right) + \beta \sum_{\tau=t}^{T-t} \delta^\tau \ln\left(\frac{c_{\tau+1}^i}{e^{k_{\tau+1}}}\right) \tag{7}$$

$$\begin{aligned} s.t \sum_{\tau=0}^{T-t} \delta^\tau c_{\tau+t}^i &\leq Rs_t + \sum_{\tau=0}^{P-t} \delta^\tau (w - d_{\tau+t}\gamma) + \delta^{P-t-1} B_l \\ &= W_t + \gamma\left(\delta^{P+1-t} \sum_{\tau=1}^{P-t-l} R_B^\tau - \sum_{\tau=l}^{P-t} \delta^\tau\right) = W_t^{B_l}. \end{aligned}$$

To consider planning problems of selves who face also the problem of selecting the starting period let us assume that SP is not started yet, it is attractive to self(t), and that self(t) is not retired.[35] Let us fix the effort cost at some commonly known level. Self(t) faces then two tasks.

[32] O'Donoghue and Rabin (2001b) defines that the agent procrastinates if his self(t) follows the strategy that involves not to complete the task even though it would be utility increasing.

[33] In this paper we do not concentrate on the 2nd order procrastination, and thus this is the reason why we do not say that the agent procrastinates if it is of the 2nd order.

[34] Since procrastination is a relevant topic to study only before retiring, our formulation for the agent's problem includes B_l. Moreover, when studying consumption-saving planning of the retirement periods, it is clear that the analysis is equivalent with the analysis of regural saving. This follows from the fact that the retired agent faces completely the same utility maximising problem as in the case of regular saving, and the maximisation problem is diffent only in the dimension of possibly different retirement time wealth.

[35] Remark that selves (i) who do not find SP attractive, (ii) who are subsequent to self(l), i.e. to self who started SP, and (iii) also selves who are already retired, i.e. self($P+1, ..., T$), do not meet the problem of starting SP. Fortunately, we have already completed the analysis about them. It has been done in the preceeding sections where SP is absent. The reasoning is following. Consider any self(t) who belongs in one of the mentioned group. Self(t)'s task is then to make the plan that maximises his intertemporal utility. Thus, the concept of his task is exactly the same with his counterpart in the world where SP is absent. Everything, except possible differences in wealth, is equivalent with the problems faced in the analysis of the basic set up.

Firstly, he must plan when to start SP if ever. Secondly, he has to make a consumption-saving plan which maximises his intertemporal utility. So far, we have fixed γ, r_B, and r, and hence we have fixed $B_j \ \forall j \in \mathbb{N}_t^P$. Self($t$) is thus able to derive all the possible considered wealth levels, $W_t^{B_j}$, and consequently, all possible intertemporal utilities he could reach if he started the attractive savings program in some particular period j. Self(t) then chooses the plan that maximises his intertemporal utility given his perceptions about future behaviour.

Since perceptions differ across the types of the agent, it is possible that not only consumption-saving plans but also planned and realised starting periods differ across the different types in the presence of SP. In the next subsections we will derive and discuss about the plans and realised actions between different selves and across the different types. Focus is mainly on when (if ever) the agent starts SP. We start by analysing NA.

4.2.1 Naive agent and SP

Consider self(t) of NA, and assume that $h^{t*} = t$. By definition, SP is attractive if starting it immediately gives higher intertemporal utility than never starting it, that is

$$\begin{aligned} & U_{t*}(C^{B_t}) - U(C^{t*}) > 0 \\ \iff \quad & \hat{U}_{t*}(C^{B_t}) - U(C^{t*}) > u(e), \end{aligned}$$

and explicitly

$$e < \left(\frac{W_t^{B_t}}{W_t} \right)^{\iota_t^{-1}} \equiv \mathcal{A}_t^{NA}, \tag{8}$$

where we have defined $\mathcal{A}_t^{NA}$ as the supremum of the effort cost to keep SP still attractive for self(t). From now on $\mathcal{A}_t^{NA}$ is called *the attractiveness.*

Procrastination is tempting if starting SP in some later period gives higher intertemporal utility than starting it immediately. Formally, procrastination is tempting if

$$\begin{aligned} \hat{U}_{t*}(C^{B_k}) - \beta\delta^{k-t}u(e) - U(C^{t*}) \quad &> \quad \hat{U}_{t*}(C^{B_t}) - u(e) - U(C^{t*}) \text{ for some } k \in \mathbb{N}_{t+1}^P \\ \Leftrightarrow e > \quad & \left(\frac{W_t^{B_t}}{W_t^{B_k}} \right)^{\frac{\iota_t^{-1}}{1-\beta\delta^{k-t}}} \quad \text{for some } k \in \mathbb{N}_{t+1}^P. \end{aligned} \tag{9}$$

Procrastination emerges then if $p^{t*} = k$, $k > t$ and self(t) anticipates that there exists self(n), $n \leq k$, such that $p^{n*} \leq k$.

The next lemma facilitates the analysis by asserting that if procrastination of l periods is lucrative so is any shorter time interval procrastination down to 1 period procrastination.

Lemma 2 *If there exists $k \in \mathbb{N}_{t+1}^{P}$ such that $e > \left(\frac{W_t^{B_t}}{W_t^{B_k}}\right)^{\frac{\iota_t^{-1}}{1-\beta\delta^{k-t}}}$, then $e > \left(\frac{W_t^{B_t}}{W_t^{B_n}}\right)^{\frac{\iota_t^{-1}}{1-\beta\delta^{n-t}}}$ for all $n \in \mathbb{N}_{t+1}^{k}$.*

Now, self(t) of NA chooses to procrastinate if the conditions (8) and (9) hold *and if* he anticipates that $\exists n \in \mathbb{N}_{t+1}^{k}$ such that self(n) chooses to start SP. Since NA considers always that his future preferences are exponential, we know by the following lemma that if self(t) anticipates that for some self(n), $n \leq p^{t*}$, the savings program is still attractive he is sure that self(n) will start the program.

Lemma 3 *If SP is attractive, exponential discounter (ED) starts it immediately.*

Along the previous lemma, we are able to say that if self(t) wants to procrastinate he never considers to procrastinate longer than one period. However, the next lemma asserts that there exists such effort cost level that self(t) procrastinates and in addition self($t+1$) will procrastinate.

Lemma 4 *Given the revised set up, $i = NA$, and a non-binding intertemporal budget constraint, there exists a level of the effort cost e such that if $p^{t*} = t+1$ then $p^{(t+1)*} = t+2$.*

As long as NA finds SP attractive and procrastination tempting he will also procrastinate. The attractiveness is, however, decreasing in time as the next lemma shows us.

Lemma 5 *Given the revised set up, and $W_t \geq e$ $\forall t \in \mathbb{N}^{P}$, the attractiveness $\mathcal{A}_t^{NA} \equiv \left(\frac{W_t^{B_t}}{W_t}\right)^{\iota_t^{-1}}$ is decreasing in t.*

From the previous lemmas it follows that NA will procrastinate until $\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right)^{\frac{\iota_t^{-1}}{1-\beta\delta}} < e < \mathcal{A}_t^{NA}$ fails to hold, and if that happens he either starts SP or abandons it for good. Which action he chooses, depends on which condition fails to hold.

Finally, from the properties of NA's reasoning and behaviour follow the next proposition.

Proposition 4 *Given the revised set up and $i = NA$, an attractive savings program will be started in the first period or it will be never started while procrastinated at least one period before abandoning it forever.*

To illustrate the import of Proposition (4) let us consider SP that could be some 401(k) program, for example, which is available to NA's self(1), but enrollment on it is not automatic. Then, the consequences might be the following. The agent finds the program worthwhile and attractive in the first period, but since there is some intermediate effort cost to enroll on the program he considers that he will start it in the next period. When the next period arrives the agent finds the program still attractive but considers that he will start it after one period. However, since our agent has fallacious perceptions about his future behaviour it happens that after some periods he does not find the program attractive any longer but considers that the effort cost is too high compared to the intertemporal utility increment from the program. Hence, he quits considering that he will start the program and abandons it forever. Furthermore, all selves who found the 401(k) program attractive but did not start it used considered wealth for optimising their consumption since they anticipated that they will start SP. The consequences of consuming considered wealth are stated in the next direct corollary of Proposition (4), and it completes the story for NA.

Corollary 1 *If NA's self(1) does not start an attractive savings program, retirement savings will be lower in the presence than in the absence of the savings program.*

Corollary (1) is one of the most important results of our procrastination analysis. When interpreted in other words it states that for an intermediate effort cost levels NA would be better off in the absence of SP than what he is in the presence of it in terms of retirement time wealth. To get the claimed result it is enough to consider the case where self(1) anticipates that he starts SP in the period $t = 2$, but self(2) does not find SP attractive any longer. In this case the first element in the self(1)'s consumption plan is $c^{1*,B_2} = \iota_1 W^{B_2}$. Now, $\iota_1 W_1^{B_2} > \iota_1 W_1 = c^{1*}$ and hence $\left(W_2 \mid c^{1*,B_2}\right) = R\left(W_1 - \iota_1 W_1^{B_2}\right) < R\left(W_1 - \iota_1 W_1\right) = \left(W_2 \mid c^{1*}\right)$. If then the effort cost is such that self(2) abandons SP, it directly follows that the rest of consumption plans are made with lower current wealth resulting in lower intertemporal utilities than in the absence of the SP. Furthermore, the retirement savings are now necessarily lower compared to the case where SP is absent.

It is a striking result that very existence of SP is enough to cause a loss in the retirement savings compared to the case where SP is absent. The results imply in addition that if the effort cost is such that the agent wants to procrastinate, it follows that the lower the effort is in this inertia range the severe the effects of procrastination will be. We know that if self(1) does not start SP, there does not exist a self who will start it. If the effort cost is then in its lowest possible inertia level, the more periods it will take before the attractiveness diminishes low enough to cause abandoning of SP. Hence, the lower is the effort cost in the inertia range the more there will be periods when selves use considered wealth resulting in even lower retirement savings.

4.2.2 Sophisticated agent and SP

We now turn to analysing sophisticated agent's behaviour. According to the definition of sophistication, SA knows his future preferences, and hence he takes SCPs into account when forming plans for consumption, saving, and about starting of SP. Just like NA also SA finds SP attractive if its immediate starting gives higher intertemporal utility than never starting it. Formally

$$\hat{U}_{t*}(C^{B_t}) - U(C^{t*}) > u(e),$$

and after some manipulation the explicit form of the attractiveness turns out to be

$$e < \frac{W_t^{B_t}}{W_t} \prod_{\tau=1}^{T-t} \left(\frac{W_{t+\tau}^{B_t}}{W_{t+\tau}} \right)^{\beta\delta^\tau} \equiv \mathcal{A}_t^{SA}$$

The attractiveness looks now slightly different and more complex at first glance than what the NA's respective condition appears to be. It is so due to SA's perfect knowledge. SA knows $\beta_{t+j} = \beta$ $\forall j \in \mathbb{N}^{T-t}$.

We are fortunately able to establish a lemma that ease the analysis by stating that the condition of attractiveness for SA's self(1) simplifies to the same form with the condition of attractiveness for NA's self(1).

Lemma 6 *Given the revised set up,* $\mathcal{A}_1^{SA} = \mathcal{A}_1^{NA}$.

The importance of Lemma (6) becomes clearer when one considers intertype comparisons. We can now always set the effort cost at any level and we know that if NA's self(1) finds SP attractive so does SA's. By Lemma (6) we, thus, avoid hesitation whether the effort cost could be at such level that SA finds SP attractive while NA does not, and vice versa.

As NA's also SA's self(t) finds procrastination tempting if from his perspective one period procrastination gives him higher intertemporal utility than starting SP immediately. Formally, procrastination is tempting for self(t) if and only if

$$e > \left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}} \prod_{\tau=1}^{T-t} \left(\frac{W_{t+\tau}^{B_t}}{W_{t+\tau}^{B_{t+1}}} \right)^{\beta\delta^{\tau}} \right)^{\frac{1}{1-\beta\delta}},$$

and hence, by using the trick shown in the proof of Lemma (6), procrastination is tempting for self(1) if

$$e > \left(\frac{W_1^{B_1}}{W_1^{B_2}} \right)^{\frac{\iota_1^{-1}}{1-\beta\delta}}. \tag{10}$$

From Condition (10) we see that procrastination is tempting for SA's self(1) whenever it is tempting for NA's self(1). It is crucial to understand at this point that even when self(1) of SA finds procrastination tempting, he knows, due to sophistication, that his future selves will make their own decisions about starting of SP, and hence self(1) does not necessarily delay starting of SP at all. On the other hand, applying the analysis from O'Donoghue and Rabin (1999c) we can say in here that SA does never start SP if and only if he prefers never starting SP to starting it immediately in period 1, i.e. SP will be never started if and only if $e \geq \mathcal{A}_1^{SA}$. It then follows that self(1) delays starting of SP only if he knows that SP will be started in some period such that a consumption structure is still utility increasing from his perspective given the realising starting period. Otherwise, he starts an attractive savings program immediately and consumes according $c_\tau^{1*} = \iota_\tau W_\tau^{B_1}$, where $\tau \in \mathbb{N}^T$. This means that self(1) resists the temptation to delay starting of SP if there does not exist a such utility increasing consumption opportunity c_1^{1*} that would trigger later selves to start SP within tolerable amount of periods given c_1^{1*}. As one might have noticed from the previous, a new problem arises for the agent. Namely, what is the optimal consumption c_1^{1*} given the period when SP is started if that particular consumption choice is chosen.

To illustrate the complexity of this new problem we consider the simplest possible case where the effort cost is at such level that delaying is tempting for self(1) and if self(1) chooses to consume according to the rule $c_1^{1*} = \iota_1 W_1^{B_2}$, then self(2) finds SP still attractive and delaying tempting but understands that if he tries to procrastinate one more period by consuming according the rule $c_2^{2*} = \iota_2 W_2^{B_3}$ SP will be unattractive for self(3). I.e. with given c_2^{2*} the $e \geq \mathcal{A}_3^{SA}$ would hold for self(3) and he would abandon the savings program for good. Hence, self(2) might deviate from consuming by the given rule and choose some strategically manipulative consumption level that makes self(3) to start SP. But then, if this consumption path induces smaller intertemporal utility to self(1) than what he would be able to get if he deviated from his given consumption rule $c_1^{1*} = \iota_1 W_1^{B_2}$ he certainly will deviate from it either by starting SP immediately or then by choosing a manipulative consumption level which induces self(2) to choose optimally from the self(1)'s viewpoint so that self(2) either starts SP or chooses a manipulative consumption level that induces self(3) to choose optimally from the viewpoints of both self(1) and self(2) etc.

As it should be clear now, there might be a continuum of possible choices for self(1) that would yield the same intertemporal utility from his perspective. To go around this problem and to implement a good device to make intertype comparisons, we will next make an assumption that constrains remarkably SA's strategy space and brings his behaviour closer to NA's behaviour. The assumption will make SA's strategy space basically the same with NA what comes to consumption possibilities with the exception that SA uses the given consumption options strategically.

Assumption 2 *Given the revised set up, $i = SA$, and an attractive non-started savings program, self(t) optimises by using the strategy that involves*
i) self(t) to choose his consumption, c_t^{t*}, from the choice set which elements are $c_t^{mc} = \iota_t W_t^{B_{t+1}}$, and $c_t^{t*,B_t} = \iota_t W_t^{B_t}$;
ii) and if self(t) chooses $c_t^{t*} = c_t^{t*,B_t}$ he starts SP immediately.

How can we interpret Assumption (2)? One possible way to consider it is to place it in the context of bounded rationality.[36] The agent knows that he can either start SP immediately, or then he can manipulate future selves' behaviour but now, due to bounded rationality, the only manipulative consumption choice self(t) is able to use in his calculations when conducting backwards induction is the consumption level that is as if he started SP in the next period.[37] Hence, we give SA

[36] Another way to consider Assumption (2) is to interpret it as a rule-of-thumb. The agent considers in every period whether to start SP or not, and if he wants to guarantee that he knows when to start it he uses the rule-of-thumb, $\iota_t W_t^{B_{t+1}}$, and knows that he the similar rule also in later periods if he decides to procrastinate then as well. More about these kind of rules of thumb see e.g. Lettau and Uhlig (1999).

[37] Note that the agent prefers as quick as possible starting of SP as long as he does not have to start it by himself. Remark then that the attractivety can be written also in the following form $\mathcal{A}_t^{SA} = \left(1 + \frac{\gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^{\tau} - \sum_{\tau=0}^{P-\tau}\delta^{\tau}\right)}{W_t}\right)^{\iota_t^{-1}}$, and hence from that expression we see that the bigger the attractivety for self(t) the more self($t-1$) consumes (W_t is the smaller the bigger is consumption in the previous period). Hence, Assumption (2) gives to the agent a weak tool to manipulate the future attractivety level in the direction that would shorten procrastination. On the other hand, it gives to the agent a powerful tool to try to make SP still attractive for a *self* in the next period. However, NA's behaviour is invariat to Assumption (2), and so, the analysis of NA stays valid also under Assumption (2).

exactly those consumption choices that NA uses when he decides whether to start SP immediately or procrastinate starting of it. We are then able to make interesting comparisons in differences between NA's and SA's behaviour when SA faces the consumption options of NA.

By Assumption (2) we can establish a simple rule for self(t) about when to start SP immediately. This rule is asserted in the following lemma, and it states that if the manipulative consumption in period t pushed the attractiveness level for self($t+1$) too low then self(t) starts the savings program immediately. The lemma certifies also that indefinite procrastination never emerges if the agent is SA.

Lemma 7 *Given the revised set up with Assumption (2), self(t) starts an attractive savings program immediately whenever* $c_t^{mc} \Rightarrow \mathcal{A}_{t+1}^{SA} \leq e$ *(or if* $t = P$*).*

Lemma (7) gives an 'end-rule' for procrastination. Intuitively, if the agent procrastinates enough it follows that the level of attractiveness diminishes low causing the correctly future behaviour predicting agent to start SP in the last period among the periods when SP is attractive. Moreover, due to our explicit choice for the manipulative consumption *the longest possible procrastination*, p_*, is exactly the same with NA. The crucial exception here is that by Lemma (7) SA's self(p_*) will always start SP while NA's counterpart still considers that he will start SP in period $p_* + 1$ but fails with his plan since actually self($p_* + 1$) abandons SP for good.[38]

In general, SA's any self(t) procrastinates only if use of the manipulative consumption leads to future behaviour that guarantees starting of SP within tolerable amount of periods from self(t)'s perspective. *The longest tolerable procrastination* for self(t), p_t^{tol}, is given by

$$\begin{aligned}
p_t^{tol} &\equiv \max\left\{p_t \in \mathbb{N}_{t+1}^{P} \mid U\left(C_t^{mc}\right) - U_{t*}\left(C^{W^{B_t}}\right) \geq 0\right\}, \text{ where} \\
U_{t*}\left(C^{mc}\right) &= \ln\left(\iota_t W_t^{B_{t+1}}\right) + \beta\left(\sum_{\tau=1}^{p_t-2} \delta^{\tau} \ln\left(\iota_{\tau+t} W_{\tau+t}^{B_{\tau+t+1}}\right) + \sum_{\tau=p_t-1}^{T-1} \delta^{\tau} \ln\left(\iota_{\tau+t} W_{\tau+t}^{B_{p_t}}\right)\right) \\
&\quad -\beta\delta^{(p_t-t-1)} \ln(e), \text{ and} \\
U_{t*}\left(C^{W^{B_t}}\right) &= \ln\left(\iota_t W_t^{B_t}\right) + \beta\sum_{\tau=1}^{T-1} \delta^{\tau} \ln\left(\iota_{\tau+t} W_{\tau+t}^{B_t}\right) - \ln(e),
\end{aligned}$$

[38] Of course, also NA's self(P) would start an attractive SP. However, since the attractiveness diminishes fairly quickly in the model when the manipulative consumption is used, and if there are many working periods as a total, the attractiveness level drops easily below 1-worthwhile level before time reaches $t = P$ causing NA to abandon SP before $t = P$.

and if $p_t^{tol} = \varnothing$ for all $p_t \in \mathbb{N}_{t+1}^P$ we say that procrastination is *intolerable* for self(t) and we denote $p_t^{tol} = t$. We then get that when SP is attractive for self(t), he first solves the longest possible procrastination, p_*, and his own tolerance, p_t^{tol}, given the rule for the manipulative consumption. Then, he solves for his future selves' tolerances, i.e. self(t) solves for p_k^{tol}, where $k \in \mathbb{N}_{t+1}^{p_*-1}$. After solving p_*, p_t^{tol}, and p_k^{tol} for $k \in \mathbb{N}_{t+1}^{p_*-1}$ self(t) is able to conduct backwards induction about his future behaviour resulting in finally the choice of starting SP or procrastinating it. These choices are fully described by the consumption choices $c_t^{t*} \in \{\iota_t W_t^{B_t}, \iota_t W_t^{B_{t+1}}\}$. Clearly, self($t$) delays starting of SP if and only if there exists self($k > t$), $k \leq p_t^{tol}$, who starts the savings program. Otherwise self(t) starts SP immediately and no procrastination emerges. The next proposition collects SA's behaviour about starting of SP in different cases.

Proposition 5 *Given the revised set up with Assumption (2), and $i = SA$, if a savings program is attractive for self(t), then*

i) self(t) procrastinates if there exists self(k), $k \in \mathbb{N}_{t+1}^{p_t^{tol}}$, who starts the savings program;

ii) self(t) starts the savings program immediately if such self(k) does not exist, or if $c_t^{mc} \Rightarrow \mathcal{A}_{t+1}^{SA} \leq e$, or if $t = P$

A straightforward application of Proposition (5) provides us the main result for the analysis of SA. This result is established in the following corollary. It asserts that if a savings program is attractive to any self(t), $t \in \mathbb{N}^P$, it will be started during the periods that belongs to self(1)'s tolerable amount of periods of procrastination. It also makes an important point of noting that as a consequence SA is able to utilise SP so that the accumulated wealth for the pension will be now greater than in the absence of SP.

Corollary 2 *Given the revised set up and $i = SA$, if a savings program is attractive for some self(t)*

i) it will be started in some period $p \in \mathbb{N}^{p_1^{tol}}$ with certainty, and thus is never abandoned for good; and

ii) the retirement savings are greater in the presence than in the absence of the savings program.

Corollary (2) distinguishes clearly between behaviour of SA and NA. While NA's self(1) starts SP only if the effort cost is so low that it is not tempting to procrastinate, self(1) of SA starts the savings program immediately also in the case where procrastination is tempting but succumbing to temptation would lead to severe procrastination from self(1)'s perspective. On the other hand, self(1) of SA may delay starting of SP when he knows that some of his subsequent incarnations within the range of tolerable procrastination will find further procrastination too costly in terms of utility, and hence starts SP. In similar situation NA's self(1) always procrastinates but his false perceptions about future behaviour cause only overconsumption, and actually, if he does not start SP immediately nor does any of his subsequent selves.

4.2.3 Learning naive agent and SP

We now turn to analysing behaviour of the learning naive agent in the presence of SP. It should be clear at this stage of the study that the logic LN uses in his decision making, excluding the first period, is the same with the logic SA uses. On the other hand, what comes to self(1) of LN the analysis is already conducted. Namely, self(1) of LN makes exactly the same choices that self(1) of NA does. Furthermore, we can maintain Assumption (2) in force also for LN. We immediately note here that if SP is attractive and procrastination is tempting for self(1) of LN he will procrastinate. This procrastination and the consumption choice based on anticipated starting of SP in the next period is clearly in line with Assumption (2). The consumption choice self(1) makes conditional on his perception about certain starting of SP in the next period is exactly the same he would choose if he procrastinated under Assumption (2). Hence, in the analysis about LN we maintain Assumption (2) with the following minor change: in the assumption, $i = SA$ is replaced with $i = LN$. We can then directly establish the first result concerning LN's behaviour.

Proposition 6 *Given the revised set up,* $i = LN$*,*
i) an attractive savings program will be started in the first period or it will be procrastinated at least one period; and
ii) if the effort cost is at such level that self(2) *of NA would abandon SP then LN's self(2) abandons SP resulting in lower level of retirement savings in the presence than in the absence of SP; otherwise*
iii) some self(t) *either starts SP or abandons it after procrastination of* $t-1$ *periods, resulting in, respectively, higher or lower retirement savings in the presesence than in the absence of SP.*

It is important to understand that the first two points in Proposition (6) follow directly from Proposition (4) and Corollary (1). Furthermore, it is a trivial consequence of our assumption about LN's ability to develop in the dimension of awareness about future SCPs. On the other hand, it emphasises the importance of time for learning about preferences: the lack of adequate time to learn about preferences can become costly. The more interesting parts of analysis consist of behaviour of those selves who have already learned something about their inconsistent planning. This is to say that we mainly concentrate on selves from the second period onwards, i.e. on the behaviour and its concequences noted in point (*iii*) of the proposition.

Once self(2) learns that he will have SCP in the future his measure for the attractiveness about SP changes from the form of NA's attractiveness measure to SA's attractiveness measure

$$\mathcal{A}_2^{LN} = \mathcal{A}_2^{SA}.$$

But then, once again we can use the analogy of the proof of Lemma (6) to establish that actually

$$\mathcal{A}_2^{LN} = \mathcal{A}_2^{NA}. \tag{11}$$

It is notable that for all LN's selves the current attractiveness level is correct but each self uses erroneous future attractiveness levels. This can be easily seen by considering self(t)'s perception about the attractiveness level for self($t+1$). Let us denote it by $\mathcal{A}_{t+1}^t$. We can then conduct the following inequality condition.

$$\mathcal{A}_{t+1}^t > \mathcal{A}_{t+1}^{LN}, \tag{12}$$

where the inequality follows, since self(t) anticipates MPC to be lower in the next period than what it in reality appears to be. Hence, in the case where self(t) finds procrastination tempting and delays starting of SP in a faith of completing the task in the next period, it can happen that when the next period arrives self($t+1$) does not consider SP attractive any longer, and hence abandons it for good. This happens whenever $\mathcal{A}_{t+1}^t > e \geq \mathcal{A}_{t+1}^{LN}$. In words, when the effort cost is in the given interval self(t) mistakenly uses the manipulative consumption since he considers that subsequent self($t+1$) will start SP. A possibility to this kind of mistake diminishes when the agent learns the size of SCP more accurately.

Since self(t) uses the biased MPC value $\iota_{(t+j)}^{(t)}$ instead of the correct ι_{t+j} for all $j \in \mathbb{N}^{T-t}$, the attractiveness level for immediately adjacent period is not the only falsely computed attractiveness level but due to miscalculated magnitudes of future manipulative consumptions, the anticipated attractiveness levels for all periods $k \geq t+1$ will be false, too.[39] LN's self(t) anticipates that if he uses the manipulative consumption in the future period $t+1$ its size will be $\iota_{(t+1)}^{(t)} W_{t+1}^{B_{t+2}}$ from which we see that $c_{t+1}^{t,mc} < c_{t+1}^{mc}$, and hence the manipulative consumption is considered to be lower than what it in reality would be if it were used in period $t+1$. This, for its part, leads to a biased estimate about the attractiveness level for self($t+1$). Biasedness for the rest of the future attractiveness levels can be naturally derived analogously.

The only difference between LN's and SA's strategy choices will follow from the biased $\beta_{(t+\tau)}^{(t)}$ that self(t) of LN uses when he chooses the strategy. Mispredicting the future attractiveness levels may cause LN to choose an suboptimal strategy. The anticipated future attractiveness levels affect directly the agent's calculations about *the anticipated longest possible procrastination* p_*^t, where t refers again to the index of the optimising self. The other source for mistakes is the anticipated future tolerance level. Self(t) considers that he will tolerate procrastination more in the future than he actually does. As we saw in the analysis of SA the longest possible procrastination and the tolerance for procrastination play leading roles in choosing the optimal strategy. Hence, extra periods of procrastination and possibly even abandoning of SP in a case where it was attractive for some self may become apparent in LN's behaviour.

Since LN's behaviour consists of parts of NA's and SA's behaviour we consider that closed form solutions about LN's behaviour would not bring any new insight to the study. Due to this fact, we complete the analysis in more interesting, descriptive, and appealing way. Namely, in the next section we will illustrate the behaviour of the different agent types both in the absence and in the presence of SP by using numerical analysis.

5 Numerical analysis

In this section we demonstrate the previous analyses by a numerical example. The example is based on calibration where $T = 50$, $P = 40$, $w = \$25,000$, $r = .0375$, $r_B = 2r$, $\beta = .85$, and $e \in [\$1.64, \$2.6]$. These choices result in $\delta = .96$ and $\gamma = \$792.74$. By using the calibration we get Figure 1 that illustrates several different aspects in four subfigures $1.1 - 1.4$. These subfigures will be discussed separately in the following subsections.

[39] Recall that the manipulative consumption for any period k was given by the rule $c_k^{k,mc} = \iota_k W_k^{B_{k+1}}$.

Figure 1.1 In Figure 1.1 we have plotted the realising starting period as a function of effort cost level. The effort cost level has been chosen to be within range $e \in [\$1.64, \$2.6]$ since when $e < \$1.64$ SP is always started immediately and on the other hand when $e > \$2.6$ SP will be abandon ed immediately. The starting period gets value 0 only if SP has been abandoned after procrastination.

As we have shown in formal analysis, NA starts the savings program immediately or then he never starts it. Due to this, we have only LN's and SA's choices in Figure 1.1. For SA it is apparent that he starts SP in the first period or then in the second period, and only if the effort cost is relatively high he waits until third period while for relatively very high effort cost he chooses to start in the second period due to the fact that waiting still one more period would cause abandoning of SP. When the effort cost is low SP stays long attractive. This feature gives more periods for LN to learn about his behaviour, and hence unnecessary abandoning due to fallacious perceptions happens only when the effort cost is relatively high or very high. Intermediate effort cost levels yield the least procrastination for LN. As a conclusion, it is clearly present in the example that SA starts always an attractive SP while LN makes mistakes when choosing the strategy and these mistakes can cause loosing SP if the effort cost level is high enough. LN never starts SP before SA but might choose a strategy that leads to abandoning it even though SA starts SP.

Figure 1.2 While the realising starting period was plotted in Figure 1.1, Figure 1.2 shows the length of procrastination in the cases where SP is finally abandoned. Length of procrastination gets value 0 when SP is abandoned without procrastination or is started after procrastination. Otherwise the length of procrastination is the last period when the agent has still considered that he will start SP in the future and hence was using the considered wealth. Since SA always starts an attractive SP and never abandons it, his choices are omitted from Figure 1.2 and only NA's and LN's choices are present.

Since the effort cost level is in such range that NA will not start SP in the first period he procrastinates as long as he finds it attractive. This means that the lower is the effort cost level the longer SP will be attarctive and hence the longer the procrastination. For LN, low level of the effort cost gives better possibility to learn his SCP and hence no abandoning with long procrastination emerges. LN makes costly mistakes only when the effort cost is high enough so that there are not enough periods to learn about SCPs before the attractiveness level decreases so low that the agent abandons SP.

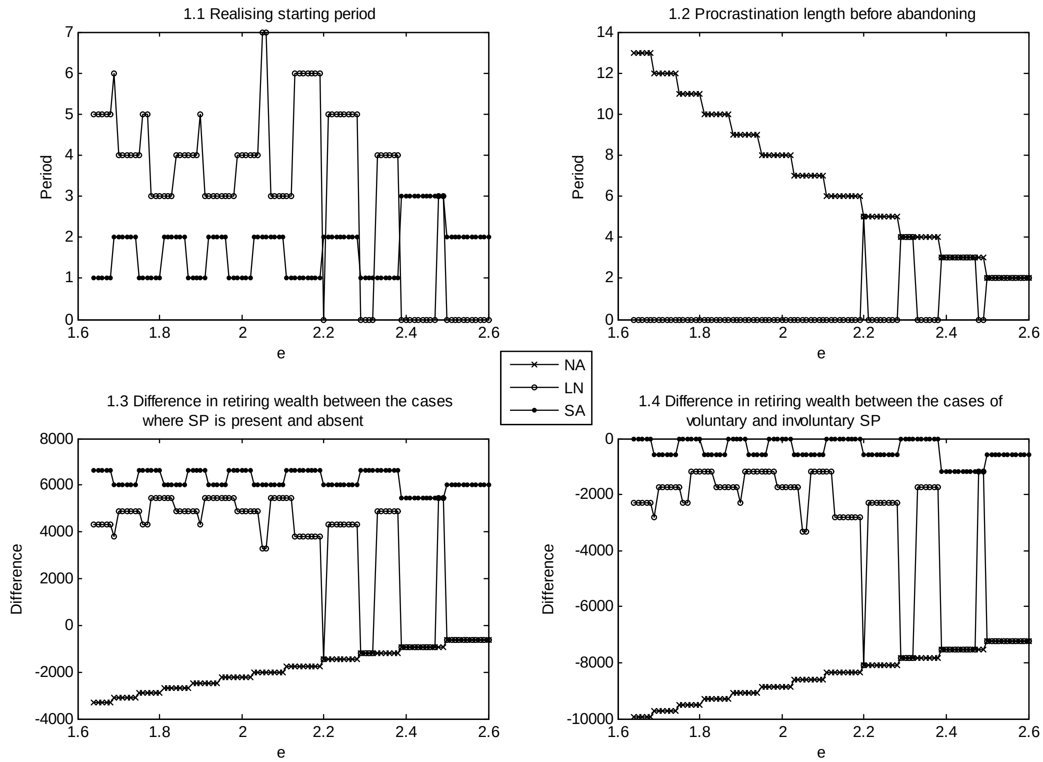

Figure 1: Results of the numerical example.

Figure 1.3 Wealth at the time of retirement between the cases where SP is present and absent is compared in Figure 1.3. Hence the figure plots the difference $W^{SP}(P+1) - W^{NSP}(P+1)$, where superscripts SP and NSP refer to existence and non-existence of SP respectively.

The biggest differences are between NA and LN as well as between NA and SA. SA always gains from the existence of SP for the given range of the effort cost while NA always suffers from it. However, the greater is the effort cost the less time NA finds SP attractive and hence the less time he then uses considered wealth without starting SP. Naturally, NA's wealth at the time of retirement in the presence of SP is the less different from the wealth at the time of retirement in the absence of SP the higher is the effort cost. LN gains almost always from the existence of SP and suffers from it only for relatively high or very high effort cost levels. For low levels of the effort cost

LN makes almost as well as SA but due the failures in planning when to start SP LN procrastinates longer than SA and has smaller wealth at the time of retirement. If the effort cost level is relatively high or very high LN does not have adequate time to learn his preferences and hence he misses SP with fallacious starting plans.

Figure 1.4 In Figure 1.4 we have compared wealth at the time of retirement in the cases of voluntary and involuntary saving programs. For a voluntary savings program we have considered SP as it is given in the revised set up while for the involuntary savings program we have considered a case where the agent is forced to save from the first period onwards according to SP.

We find that involuntary enrollment in the program would result always higher or the same wealth at the time of retirement than voluntary enrollment, which is not surprising since with respect to the magnitude of retirement wealth the agent can't make better off than what he is able to attain if he started SP in the first period. For SA it is almost the same whether SP is voluntary or involutary, he makes almost as well in both cases. For LN the same holds for low levels of effort cost but when the effort cost is relatively high or very high he starts to suffer from the voluntary system. Clearly, for NA the involuntary system would make the greatest improvement as he would gain from it relatively a great deal of money.

6 Conclusions

In this paper we focused on some existence effects of a retirement savings program. The analysis of the basic set up reveals us that when the agent has hyperbolic intertemporal preferences with logarithmic instantaneous utility function backward looking behaviour does not affect the realising saving choices when retirement savings program is absent. All the agent types, naive, learning naive, and sophisticate, make the same consumption choices but only the sophisticated agent is able to form a time-consistent consumption plan. Hence, the retirement savings are, as well, the same across the different types in the absence of the savings program. When a lucrative retirement savings program with fixed contribution is incorporated into the basic set up the agent type matters and retirement savings can be different between the types. Enrollment on the savings program is considered costly in terms of utility while the retirement savings program is made lucrative by putting higher interest rate on savings through the program than what it is on the regular checking account savings.

The most striking result is that for certain substatially low levels of the effort cost the plain existence of the retirement savings program yields always lower retirement savings for the naive agent and some times also for the learning naive agent than what they would be able to attain in the absence of the retirement savings program. For the same levels of the effort cost, the sophisticated agent, for his part, always gains from the existence of the retirement savings program and attains higher retirement savings through the program than what he would otherwise able to get. The results follow from the effort cost, anticipation about future behaviour and flexibility in the budget constraint. The effort cost on enrolling on the retirement savings program makes the agents to procrastinate their completion of the starting task in a faith that there exists a self in the future who will start the savings program. Since the agent believes that he really is going to start the program in the future he takes the increased income stream into account and consumes some fractions from that even before the enrollment. This is possible if there is flexibility in the budget constraint. Finally, if the anticipation about future behaviour happens to be incorrect in a sense that future selves do not find the program attaractive any longer, it will be abandoned. But then, due to preceding selves' consumption of the imaginary wealth, the retirement savings can not be as high as what they would have been if the savings program were completely absent.

The naive agent always considers that his subsequent self will start the program, and actually in our model it happens then that if the naive agent does not start the savings program in the very first period it will be never started but procrastinated at least one period, resulting in decrease in retirement savings compared to the case where the savings program is absent. The sophisticated agent is the opposite to the naive and always starts an attractive savings program in some period that ensures higher retirement savings than in the absence of the program. This is possible for selves of the sophisticated agent since they know perfectly their future behaviour and are thus able to choose strategically correct actions to make it sure that an attractive retirement savings program is started in some period. In some cases, learning naive gains from the existence of the savings program and for some cases he suffers from it. All that matters for the learning naive is the amount of the periods he has for learning about his self-control problems. That is, the effort cost must be on such level that he has enough periods to learn before the attractivety drops too low. If the latter happened the retirement savings would be lower than in the absence of retirement savings program and if the former happened then the case would be opposite for retirement savings.

In the numerical analysis we illustrated also the difference in retirement savings between the cases of non-automatic enrollment and automatic enrollment on a retirement savings program. It is obvious that in our set up the automatic enrollment yields greater retirement savings than the non-automatic does. Also, the automatic enrollment always yields higher retirement savings than what the savings are in the absence of the program. However, an important finding from this is that once the retirement savings program is implemented and while the difference in retirement savings is fairly small for learning naives and sophisticates, naives would gain relatively much from automatic enrollment. This implies that sophisticates do not suffer the chosen system never alot but for naives the type of a chosen system can have a big difference. Whether the chosen system affects learning naives depends the most on the magnitude of the effort cost. If the effort level is relatively low they have enough time to learn and hence they are more or less the same in the cases of involuntary and voluntary systems while for high and very high levels of the effort cost involuntary system would be definitely better for them.

Finally, since there is a growing tendency to put more responsibility and pressure on individuals about their own retirement funding, we should concentrate on implications of a retirement savings program and how the enrollment on it should be done. As it is now clear, the results establish that fairly low levels of the effort cost can make relatively big differences in retirement wealth. If we believed that only naive types exist then, in principle, all the saving possibilities should be equal so that no optional saving programs existed, or the effort cost on enrolling on retirement saving programs should be very low, or the involuntary system should be exercised. On the other hand, if we believed that there exist only sophisticates the only thing to be cared of would be to keep the effort cost only on intermediately low so that the savings program remains attarctive for an individual who is just on the phase of starting to work regularly. We consider, however, that in reality neither of these extreme types probably does exist and all individuals are more or less equiped with capabilities to learn about them selves what comes to retirement saving, and that the capability to learn just varies across individuals.[40] In addition, since we do not know the fractions of these different types in economy, then the effort cost should be lowered below the level that makes procrastination tempting. As easy as it sounds there exists a problem, though. It is reasonable to believe that the lower we would like to push the effort cost level the more expensive it would be.

[40] Then, one could consider naives as extremely slow to learn and sophisticates extremely quick to learn.

Thus, if it is too costly to push the effort cost level low enough, to avoid losses in naives' retirement savings we should let the effort cost be high. But then, also learning naives will start to suffer from the existence of the savings program since they do not have enough time to learn from their self-control problems. Hence the main problem with the existence of the savings program is that if the effort cost is on low level but procrastination is tempting, naives are hurt while learning naives and sophisticates are doing quite well with their savings. Letting the effort cost to rise helps naives but hurts then learning naives, while sophisticates are still doing well. If the effort cost is still let to be higher both naives and learning naives will be out of interest to save along retirement savings program and they do not suffer that much about the existence of SP, while sophisticates still do well. This implies a meta-result which says that if we consider all the agents to be learning naives with different capabilities to learn, then if lowering the effort cost is expensive relative to hastening the learning capability, possibilities to learn about self-control problems should be improved. If them both are expensive, it should be tried to avoid heterogeneity in different possibilities to save or let the effort cost to rise so that only sophisticates start it while naives and learning naives abandon it fairly quickly. Lastly, the results impy intuitively that if the publicly financed retirement system provided originally more than the savings program yields for SA, the publicly financed system could be maintained by decreasing the replacement rate on the respective of SAs ability to save along SP and all the types should accepted it.

References

Ainslie, G. (1992): *Picoeconomics: The Strategic Interaction of Successive Motivational States with the Person*. Cambridge University Press.

Angletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001): "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives*, 15, 47–68.

Börsch-Supan, A., and A. Brugiavini (2001): "Savings: The Policy Debate in Europe," *Oxford Review of Economic Policy*, 17, 116–143.

Choi, J., D. Laibson, B. Madrian, and A. Metrick (2004a): "For Better or For Worse: Default Effects and 401(K) Savings Behaviour," in *Perspectives in the Economics of Aging*, ed. by D. A. Wise, pp. 81–121, Chicago, IL. University of Chicago Press.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004b): "Optimal Defaults and Active Decisions," *Working Paper.*

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351–401.

Harris, C., and D. Laibson (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 69, 935–957.

Holzmann, R., M. Orenstein, and M. Rutkowski (eds.) (2003): *Pension Reform in Europe: Process and Progress*. World Bank, Washington D.C.

Hsieh, C.-T. (2003): "Do Consumers React to Anticipated Income Changes? Evidence from Alaska Permanent Fund," *American Economic Review*, 93, 397–405.

Laibson, D. (1994): "Essays in Hyperbolic Discounting," *Ph.D. Dissertation, MIT.*

——— (1996): "Hyperbolic Discount Functions, Undersaving, and Savings Policy," *NBER Working Paper Series*, WP 5635.

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

Laibson, D., A. Repetto, and J. Tobacman (1998): "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity*, 1998, 91–172.

——— (2003): "A Debt Puzzle," in *Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, pp. 228–266, Princeton. Princeton University Press.

Lettau, M., and H. Uhlig (1999): "Rules of Thumb versus Dynamic Programming," *American Economic Review*, 89, 148–174.

Loewenstein, G., and D. Prelec (1992): "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107, 573–597.

Madrian, B. C., and D. F. Shea (2001): "The Power of Suggestion: Inertia in 401(K) Participation and Savings Behavior," *Quarterly Journal of Economics*, CXVI, 1149–1187.

Miravete, E. J. (2003): "Choosing the Wrong Calling Plan? Ignorance and Learning," *American Economic Review*, 93, 297–310.

O'Donoghue, T., and M. Rabin (1999a): "Doing It Now or Later," *The American Economic Review*, 89, 103–124.

O'Donoghue, T., and M. Rabin (1999b): "Incentives for Procrastinators," *Quarterly Journal of Economics*, 114, 769–816.

——— (1999c): "Procrastination in Preparing for Retirement," in *Behavioral Dimensions of Retirement Economics*, ed. by H. Aaron. The Brookings Institution.

——— (2001a): "Addiction and Present-Biased Preferences," *Unpublished manuscript.*

——— (2001b): "Choice and Procrastination," *Quarterly Journal of Economics*, 116(1), 121–160.

——— (forthcoming): "Self Awareness and Self Control," in *Now or Later: Economic and Psychological Perspectives on Intertemporal Choice*, ed. by R. Baumeister, G. Loewenstein, and D. Read. Russel Sage Foundation Press.

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

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

Stortz, R. H. (1955-1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23, 165–180.

Appendix

A Different agent types

Let us now present the definition of different types of the agent in terms of agent's awareness about his *future* SCP. Our type classification follows the lines of the verbal type classification given in O'Donoghue and Rabin (2001b, forthcoming). However, we depart from their definition by defining a new agent type, the learning naive agent (*LN)*, which replaces the partially naive agent (*PN)*.[41]

By definition the agent has SCP if he uses non-exponential discounting. The agent can be aware of his future SCPs only if he is aware of the fact that he will use non-exponential discounting in the subsequent periods as well.[42] Since β is the only parameter that causes discounting to be non-exponential in our model, the definition of awareness about future SCPs reduces to defining agent's self(t)'s perceptions about $\beta_{(t+j)}$ We can then define awareness formally in the following way.

Definition 5 *Self(t) is aware of his future SCPs if* $\beta^{(t)}_{(t+j)} \neq 1$ *for* $j \in \mathbb{N}^{T-t}$, *where* $\beta^{(t)}_{(t+j)}$ *denotes self*(t)*'s perception about* β_{t+j}, *i.e. a perception about the value of* β *in self(t+j)'s utility function* U_{t+j}.

Definition (5) defines only whether self(t) is aware of the fact that $\beta \neq 1$ also for the future selves. However, the definition does not tell us anything about agent's knowledge of β's true value in the later periods. It is now completely possible that the agent is aware of the fact $\beta_{t+j} \neq 1$ but $\beta^{(t)}_{(t+j)} \neq \beta$. It then follows that the definition in question does not tell us whether the agent knows his subsequent selves' real preferences even when he is aware about future SCPs. To distinguish between the different agent types we use agent's self(t)*'s* different levels of awareness about $\beta_{(t+j)}$.

Starting from *NA,* we say that the agent is *NA* if his every self*(*$t \in \mathbb{N}^{T-2}$*)* is unaware about future SCPs. According to Definition (5), self(t) is unaware about future SCPs if $\beta^{(t)}_{(t+j)} = 1 \ \forall j \in \mathbb{N}^{T-t}$.[43] In the same manner we define *SA*. We say that the agent is *SA* if his every self*(*$t \in \mathbb{N}^{T-2}$*)* is *perfectly* aware about future SCPs. Then, by Definition (5) self(t) is aware about future SCPs if $\beta^{(t)}_{(t+j)} \neq 1 \ \forall j \in \mathbb{N}^{T-t}$. To make the agent to be perfectly aware about future SCPs we require $\beta^{(t)}_{(t+j)} = \beta \ \forall j \in \mathbb{N}^{T-t}$.[44]

There is still one undefined type, namely *LN.* The motivation for introducing the learning naive agent into the type category of hyperbolic agents is, in the first place, that in the recent literature

[41] Along O'Donoghue and Rabin (2001b) *PN* is aware of his future SCP but he constantly underestimates its magnitude. In other words, any *self(t)* of *PN* considers that he will have SCP in the future with the magnitude $\hat{\beta}$, but he considers that $1 > \hat{\beta} > \beta$, and hence is over-optimistic about his capabilities to control himself in the future. Recall, the bigger β is the smaller SCP is.

[42] Here we implicitely assume that the agent understands what it means if his intertemporal utility function incorporates non-exponential discounting.

[43] Only selves from 1 to $T-2$ are required to be unaware, since there is no meaning to discuss about awareness in the trivial periods $T-1$ and T. For *self(* $T-1$ *)* there is only passive *self(T)* left who does not have the intertemporal preferences. Hence, it does not matter what *self(* $T-1$ *)* considers about the intertemporal preferences of *self(T)*. Moreover, since by the construction of our model we assume that all the types of the agent know perfectly their last day in live, we can set with out loss of generality that if $t = T$ then $j = 0$ and consequently $\beta^{(T)}_{(T+j)} = \beta$ for all $i \in \mathcal{I}$.

[44] Note that perfect awareness implies awareness. If the $\beta^{(t)}_{(t+j)} = \beta \ \forall j \in \mathbb{N}^{T-t}$ and $\beta \neq 1$, it is clear that $\left(\beta^{(t)}_{(t+j)} = \beta \ \forall j \in \mathbb{N}^{T-t}\right) \Rightarrow \left(\beta^{(t)}_{(t+j)} \neq 1 \ \forall j \in \mathbb{N}^{T-t}\right)$.

a type of the agent is assumed to be in a status quo. If the agent is classified to be *NA* he will be *NA* also after infinite amount of periods. Same holds for *PN* and *SA*. More importantly, they do not have any possibilities to learn about their past behaviour. However, when time elapses agent's current information changes if he has any kind of recall, and hence agent's possibilities of understanding that he has future SCPs should be improved.[45] So, it is completely possible that the agent learns about his preferences and thus about his SCP.[46] In our analysis *LN* represents this kind of learning type of the agent.

We say that the agent is *LN* if he is like *NA* in the first period but in subsequent periods he becomes more and more aware about future SCPs' true magnitude. Before giving an exact formal definition we describe the properties of *LN*. We consider *LN* is naive in the sense that in the first period he is not aware of future SCPs, thus $\beta^{(1)}_{(1+j)} = 1 \; \forall j \in \mathbb{N}^{T-1}$. This assumption can be motivated, for example, with the following reasoning. In the first period the agent does not have any past plans which could provide him information about his intertemporal preferences in different periods. Hence, the agent considers that his preferences will be the same in the future as they are in the first period. We consider *LN* is learning in the sense that in the second period self(2) understands that his intertemporal preferences differ from the preferences he thought in the previous period to have in the current period. Hence, he learns he has SCP. To estimate the size of future SCPs we assume that *LN* uses the simple rule $\beta^{(t)}_{(t+j)} = \frac{\beta_t + \beta^{(t-1)}_{(t)}}{2}$, where β_t is the real β in period t and $\beta^{(t-1)}_{(t)}$ is the agent's estimate in the previous period for the current period's β.[47] It is easy to check that *LN* is always over-optimistic about his future behaviour, i.e. $\beta < \beta^{(t)}_{(t+j)} \leq 1$ for all $t \in \mathbb{N}^{T-t}$, and thus he considers that his SCP will be smaller in the future than what it is at the present moment. However, due to *LN*'s backward-looking behaviour his estimates get closer and closer to the true β. Finally, if T is large enough, i.e. if there are enough periods to attain high convergence towards β, *LN*'s perceptions about $\beta_{(t+j)}$ are almost the same with *SA*'s perceptions that are the perfect ones. Putting it loosely, *LN*'s perceptions about his future preferences develop from *NA*'s to almost *SA*'s perceptions during his life time. This is due to his capability of learning, which is characterised with perceptions that satisfies $\beta^{(1)}_{(1+j)} = 1 \; \forall j \in \mathbb{N}^{T-1}$ and $\beta < \beta^{(t+1)}_{(t+j+1)} < \beta^{(t)}_{(t+j)} < 1 \; \forall t, j \in \mathbb{N}^{T-2}_{2}$.

To this end, we can express our formal definition of different types of the agent.

Definition 6 *Let $\beta^{(t)}_{(t+j)}$ denote self(t)'s perceptions about β_{t+j}, and let the true state be such that $\beta_{t+j} = \beta_t = \beta < 1 \, \forall t \in \mathbb{N}^T$ and $\forall \; j \in \mathbb{N}^{T-t}$. Then*
a) The agent is Naive iff $\beta^{(t)}_{(t+j)} = 1 \; \forall t \in \mathbb{N}^{T-2}, \forall j \in \mathbb{N}^{T-t}$;
b) The agent is Learning naive iff $\beta^{(1)}_{(1+j)} = 1 \; \forall j \in \mathbb{N}^{T-1}$ and $\beta < \beta^{(t+1)}_{(t+j+1)} < \beta^{(t)}_{(t+j)} < 1 \; \forall t \in$

[45] See e.g. Miravete (2003).
[46] However, the fact that the agent had SCP in the preceding periods does not necessarily imply that he could or would not think he does not have them in the future. This topic is problematic, since there is evidence in the literature that even old people suffer from SCPs, and on the other hand there is evidence that even animals learn from their past behaviour, hence why people should not recall and then learn.
[47] *Self(t)* has to always know β_t since otherwise the agent would not know even his current intertemporal preferences. Note also, that if *self(t)* used the simplest possible rule $\beta^{(t)}_{(t+j)} = \beta_t$ his perceptions would be perfect and he would be immediately *SA* after the first period. So, in some sense the useage of the simplest rule requires then more sophistication than using of less simple ones. Naturally, these types of sophistication are in different dimensions and not under contemplation in here.

$\mathbb{N}_2^{T-2}, \forall j \in \mathbb{N}^{T-t}$;

c) The agent is Sophisticated iff $\beta^{(t)}_{(t+j)} = \beta\ \forall t \in \mathbb{N}^{T-2}, \forall j \in \mathbb{N}^{T-t}$.

In the recent literature a terminology between rationality and irrationality with respect to a level of sophistication is very confusing, sometimes even mixed up and erroneous. We note that according to this literature *NA* in Definition (6) could be called also as the irrational agent. We refrain using this term since choices that *NA* makes are all completely rational from each self's perspective. By respecting these manners we do not call *SA* as the rational agent, even though *SA* has rational expectations and a perfect foresight. About *LN* we are still eager to remark that by the definition the agent is *LN* whenever his perceptions about future develop to correct direction, no matter how this learning is gained.

B Proofs

B.1 Proof of Proposition (1)

Proof. *NA* Consider now NA's consumption problem. Pick an arbitrary self(t) of NA. Since $i = NA$ then by Definition (1) self(t) has perceptions $\beta^{(t)}_{(t+j)} = 1\ \forall j \in \mathbb{N}^{T-t}$. For any given consumption history of self(t)'s past selves, self(t) considers that his maximisation problem is given as follows.

$$\underset{\{c^t_{t+j}\}_{j=0}^{T-t}}{Max} \{\ln c^t_t + \beta \sum_{\tau=1} \delta^\tau \ln c^t_{t+\tau}\}$$
$$s.t.\ c^t_t + \sum_{\tau=1}^{T-t} \delta^\tau c^t_{\tau+t} \leq Rs_t + w \sum_{\tau=0}^{P-t} \delta^\tau.$$

Taking derivatives with respect to consumptions yield the following first order conditions

$$c^t_t = \beta^{-1} c^t_{t+1} = \beta^{-1} c^t_{t+j}, \forall j \in \mathbb{N}^{T-t}.$$

Substituting $c^t_{t+j} = \beta c^t_t$ into the budget constraint results in the following optimal consumption for the period t.

$$c^{t*}_t = \frac{Rs_t + w\sum_{\tau=0}^{P-t}\delta^\tau}{1+\beta\sum_{\tau=1}^{T-t}\delta^\tau} = \frac{W_t}{1+\beta\sum_{\tau=1}^{T-t}\delta^\tau} = \iota_t W_t, \tag{13}$$

where $\iota_t = (1+\beta\sum_{\tau=1}^{T-t}\delta^\tau)^{-1} = \frac{1-\delta}{1-\delta+\beta(1-\delta^{T-t})}$. Now, $\iota_t = \frac{\partial c_t}{\partial W_t}$, and hence ι_t is self(t)'s marginal propensity to consume (MPC). Clearly, ι_t is not constant in time but monotonic and increases in t. Since $\beta^{(t)}_{(t+j)} = 1\ \forall j \in \mathbb{N}^{T-t}$, self($t$) has an anticipation $\iota_{t+j} = \left(1+\sum_{\tau=1}^{T-t-j}\delta^\tau\right)^{-1}$ for $\forall j \in \mathbb{N}^{T-t}$. Thus, self($t$) forms his consumption plan for all the subsequent periods by using Eq.(13) and the following rule from the first order condition

$$c^{t*}_{t+j} = \beta c^{t*}_t. \tag{14}$$

Together those imply

$$c^{t*}_{t+j} = \beta \iota_t W_t\ \forall j \in \mathbb{N}^{T-t}.$$

Since self(t) was arbitrarily chosen, Eqs.(13) and (14) characterise the consumption plan for any self($t \in N^T$), which is in general form

$$C_{NA}^{t*} = \{\iota_t W_t, \underbrace{\beta\iota_t W_t, \beta\iota_t W_t, ..., \beta\iota_t W_t}_{T-t \text{ elements}}\}.$$

Finally, each self completes the maximisation in his turn, and hence a sequence of realised consumption choices from period t onwards is given by

$$C_{NA}^{t} = \{\iota_j W_j\}_{j=t}^{T}.$$

From Eqs.(13) and (14) one can easily derive

$$\begin{aligned}
c_t^{t*} &= \beta^{-1} c_{(t+1)}^{t*} = \beta^{-1} c_{(t+j)}^{t*} = \beta^{-1} c_{(t+j+k)}^{t*} \Rightarrow c_{(t+j)}^{t*} = c_{(t+j+k)}^{t*} \\
c_{t+j}^{(t+j)*} &= \beta^{-1} c_{t+j+1}^{(t+j)*} = \beta^{-1} c_{t+j+k}^{(t+j)*} \Rightarrow c_{t+j}^{(t+j)*} = \beta^{-1} c_{t+j+k}^{(t+j)*} \\
c_{t+j}^{t*} &= c_{t+j+k}^{t*} \neq \beta^{-1} c_{t+j+k}^{(t+j)*} = c_{t+j}^{(t+j)*} \\
&\Rightarrow c_{t+j}^{t*} \neq c_{t+j}^{(t+j)*} \ \forall t \in \mathbb{N}^{T-2} \text{ and } \forall j \in \mathbb{N}^{T-t-1}.
\end{aligned} \tag{15}$$

Now, the last line implies that $C_{NA}^{t} \neq C_{NA}^{t*}$ for all $t \in \mathbb{N}^T \setminus \{T-1, T\}$. The last and the second to last periods are left out as trivial ones.

SA Consider now any (non-trivial) self(t) of *SA* and his consumption problem. By Definition (6) every *SA*'s self(t) has perceptions $\beta_{(t+j)}^{(t)} = \beta \ \forall j \in \mathbb{N}^{T-t}$ implying that every self(t) knows perfectly future incarnations' intertemporal preferences. Then, self(t) makes his consumption-saving plan conditional on known intertemporal preferences of his future selves. That is, self(t) chooses such consumption choice for period t that maximises his intertemporal utility given the reaction of future selves to his current consumption choice.

To find out the best consumption-saving choice for period t, given that future selves stick to their optimal consumption choices, self(t) uses backwards induction and solves the Bellman equation.[48] Self(t)'s strategy is, thus, a mapping from the state variables to the choice variable. To obtain an equilibrium, any self(t)'s strategy has to be such that it is optimal given all other selves', self$(k \in \mathbb{N}^T \setminus \{t\})$, strategies. Then clearly, the equilibrium is a fixed point in a strategy space.

Take self(t)'s perspective. His value function, V_{t+j}^t, for any future period $t+j$, satisfies the following recursion

$$V_{t+j}^t(W_{t+j}) = u(c_{t+j}^{(t+j)*}) + \delta V_{t+j+1}^t(R(W_{t+j} - c_{t+j}^{(t+j)*})), \tag{16}$$

where the optimal consumption $c_{t+j}^{(t+j)*}$ chosen by self*(t+j)* is a function of wealth W_{t+j}, and the discount factor is δ, since self(t) uses exponential discounting from period $t+1$ onwards. In period t, self(t) discounts future utility with the discount factor $\beta\delta$, and hence the value function for period t is

$$V_t^t(W_t) = u(c_t^{t*}) + \beta\delta V_{t+1}^t(R(W_t - c_t^{t*})).$$

[48] Our derivation of *SA*'s behaviour follows Laibson (1997), Harris and Laibson (2001) Laibson et al. (2003).

Now, to find the optimal consumption choice c_t^{t*} which solves the value function above, self(t) solves

$$c_t^{t*} = \max_{c_t^t} \left\{ u(c_t^t) + \beta\delta V_{t+1}^t(R(W_t - c_t^t)) \right\}.$$

Let us then consider the explicit solution for self(t)'s consumption problem. To use backwards induction, we start from the last period. From the assumption of '*no bequest motive*' follows $\iota_T = 1$, hence

$$s_{T+1} = 0 \Rightarrow c_T^{T*} = \iota_T W_T = W_T \Rightarrow V_T^t(W_T) = \ln W_T.$$

Then, self(t) solves rest of the value functions recursively by using Eq.(16). Thus, for the second to last period the value function is of the form

$$\begin{aligned} V_{T-1}^t(W_{T-1}) &= \ln c_{T-1}^{(T-1)*} + \delta \ln(R(W_{T-1} - c_{T-1}^{(T-1)*})) \\ &= \ln \iota_{T-1} W_{T-1} + \delta \ln(R(W_{T-1} - \iota_{T-1} W_{T-1})) \\ &= \ln W_{T-1} + \delta \ln(R(1 - \iota_{T-1}) W_{T-1}) + \ln \iota_{T-1} \\ &= (1+\delta) \ln W_{T-1} + Z_{T-1}, \end{aligned}$$

where $Z_{T-1} = \ln \iota_{T-1} + \delta \ln(R(1-\iota_{T-1}))$ is constant with respect to choice variable.[49] From self(t)'s viewpoint the value functions for periods $t+j$ are generally of the form.

$$V_{t+j}^t = \left(\sum_{\tau=0}^{T-(t+j)} \delta^\tau \right) \ln W_{t+j} + Z_{t+j},$$

where the "periodical" constant Z_{t+j} is of the form

$$Z_{t+j} = \ln \iota_{t+j} + \sum_{\tau=1}^{T-(t+j)} \delta^\tau \left(\ln R(1 - \iota_{t+j}) \right).$$

Particularly, for period t the value function is

$$V_t^t(W_t) = \max_{c_t^t} \left\{ \ln c_t^t + \beta\delta \left(\left(\sum_{\tau=0}^{T-(t+1)} \delta^\tau \right) \ln(R(W_t - c_t^t)) + Z_{t+j} \right) \right\}.$$

Let us then solve the maximisation problem above. Taking the derivative with respect to c_t^t and setting it equal to zero yields

$$\frac{1}{c_t^t} - \frac{\beta \sum_{\tau=1}^{T-t} \delta^\tau}{W_t - c_t^t} = 0.$$

Solving c_t^t yields

$$c_t^{t*} = \frac{W_t}{1 + \beta \sum_{\tau=1}^{T-t} \delta^\tau} = \iota_t W_t. \tag{17}$$

[49] The expression $Z_{T-1} = \ln \iota_{T-1} + \delta \ln(R(1 - \iota_{T-1}))$ is obviously constant, since the marginal propensity to consume ι_{T-1}, gross interest rate R, and the discount factor δ are all independent of the choice variable.

Since we chose t arbitrarily, the optimising rule (17) gives the optimal current consumption for any self. Self(t) uses Eq.(17) to derive his optimal current consumption, and hence after solving this, his consumption plan is complete.[50] The complete consumption plan for any self(t) is then

$$C_{SA}^{t*} = \{\iota_j W_j\}_{j=t}^{T} = \left\{ \frac{Rs_j + w_j \sum_{\tau=0}^{P-j} \delta^\tau}{1 + \beta \sum_{\tau=1}^{T-j} \delta^\tau} \right\}_{j=t}^{T}. \tag{18}$$

C_{SA}^{t*} is also the path of realised consumptions, i.e. $C_{SA}^{t*} = C_{SA}^{t}$ for all $t \in \mathbb{N}^T$ since each self uses Eq.(17) to solve optimal current consumption.

LN To derive *LN*'s consumption-saving behaviour we start the analysis from self(1) of *LN*. We know that for *LN*'s self(1) holds $\beta_{(1+j)}^{(1)} = 1$, hence his consumption plan coincides with *NA*'s self(1). Thus,

$$C_{LN}^{1*} = C_{NA}^{1*} = \{c_1^{1*,NA}, \beta c_1^{1*,NA}, ..., \beta c_1^{1*,NA}\},$$

where the number of elements in the subsequence $\{\beta c_1^{1*,NA}, ..., \beta c_1^{1*,NA}\}$ is naturally $T-1$. Self(2) recognises $c_2^{2*} \neq c_2^{1*}$, and learns that he has SCP.[51] Self(2) then estimates the size of *future* SCP for periods $t = 2 + j$ by the given rule, i.e. $\beta_{(2+j)}^{(2)} = \frac{\beta + \beta_{(1+j)}^{(1)}}{2} > \beta$. Then, consumption plan C_{LN}^{2*} is formed similarly with *SA*. *LN* behaves now exactly like *SA* but uses upwards biased estimates about his future SCP. C_{LN}^{2*} is then of the form

$$C_{LN}^{2*} = \left\{ \frac{Rs_2 + w \sum_{\tau=0}^{P-2} \delta^\tau}{1 + \beta \sum_{\tau=1}^{T-2} \delta^\tau}, \left(\frac{Rs_j + w_j \sum_{\tau=0}^{P-j} \delta^\tau}{1 + \beta_{(j)}^{(2)} \sum_{\tau=1}^{T-j} \delta^\tau} \right)_{j=3}^{T} \right\}.$$

In general, for any t, C_{LN}^{t*} is then of the form

$$\begin{aligned} C_{LN}^{t*} &= \left\{ \frac{Rs_t + w \sum_{\tau=0}^{P-t} \delta^\tau}{1 + \beta \sum_{\tau=1}^{T-t} \delta^\tau}, \left(\frac{Rs_j + w_j \sum_{\tau=0}^{P-j} \delta^\tau}{1 + \beta_{(j)}^{(t)} \sum_{\tau=1}^{T-j} \delta^\tau} \right)_{j=t+1}^{T} \right\} \\ &= \left\{ \iota_t W_t, \left(\iota_{(j)}^{(t)} W_j^{t*} \right)_{j=t+1}^{T} \right\}, \end{aligned}$$

where $\iota_{(j)}^{(t)}$ denotes self(t)'s perception about self(*j*)'s MPC and W_j^{t*} is the anticipated wealth for period j. Since the first element in consumption plan is always the realing consumption, the sequence of realised consumption choices for *LN* will be

$$C_{LN}^{t} = \left\{ \frac{Rs_j + w_j \sum_{\tau=0}^{P-j} \delta^\tau}{1 + \beta \sum_{\tau=1}^{T-j} \delta^\tau} \right\}_{j=t}^{T} = \{\iota_j W_j\}_{j=t}^{T},$$

[50] We remark that self(t)'s only task is to plan his current consumption, since the rest of the consumption decicions are made by the future selves and self(t) takes them as given.

[51] It is important to notice that according to the analysis of *SA*'s behaviour, no matter how self(2) forms his consumption plan, i.e. what $\beta_{(t+j)}^{(t)}$ he uses, he will always find $c_2^{2*} \neq c_2^{1*}$ when self(1) has formed his plan as *NA*.

for any $t \in \mathbb{N}^T$.

To complete the proof we note that now $C^t_{NA} = C^t_{SA} = C^t_{LN}$. ■

B.2 Proof of Lemma (1)

Proof. Pick an arbitrary $t \in \mathbb{N}^T \setminus \{T\}$. Rewrite the claim by using $c_t = \iota_t W_t$, $c_{t+1} = \iota_{t+1} W_{t+1}$, then show that

$$\iota_t W_t - \iota_{t+1} W_{t+1} > 0,$$

which is equivalent with showing

$$\iota_t W_t - \iota_{t+1} W_t R(1-\iota_t) > 0.$$

Solving ι_t gives inequality

$$\iota_t > \frac{R\iota_{t+1}}{1+R\iota_{t+1}}. \tag{19}$$

Contemplate then r.h.s of reproduced claim (19) and denote $\frac{R\iota_{t+1}}{1+R\iota_{t+1}} \equiv H$. Then, elaborating H results in

$$\begin{aligned} H \quad &= \quad \frac{1}{\frac{1}{R\iota_{t+1}}+1} = \frac{1}{R^{-1}(\iota_{t+1})^{-1}+1} \overset{(R^{-1}=\delta)}{=} \frac{1}{\delta(1+\beta\sum_{\tau=1}^{T-t-1}\delta^{\tau})+1} \\ &\overset{(R^{-1}=\delta)}{=} \frac{1}{\delta(1+\beta\sum_{\tau=1}^{T-t-1}\delta^{\tau})+1} = \frac{1}{(\delta+\beta)+\left(1+\beta\sum_{\tau=1}^{T-t}\delta^{\tau}\right)} \end{aligned}$$

Now, we have

$$\iota_t = \frac{1}{\left(1+\beta\sum_{\tau=1}^{T-t}\delta^{\tau}\right)} > \frac{1}{(\delta+\beta)+\left(1+\beta\sum_{\tau=1}^{T-t}\delta^{\tau}\right)} = H,$$

which completes the proof since t was arbitrarily chosen. ■

B.3 Proof of Lemma (2)

Proof. (i)-(ii) Follows directly from Proposition (1).

iii) Consider now any $t < P$ and the consumption plans C^{t*}_{NA} and $C^{(t+1)*}_{NA}$. Since $C^{t*}_{NA} \neq C^t_{NA}$ and particularly $c^{t*}_{t+1} < c_{t+1}$, we know that $W^{t*}_{t+j+1} > W_{t+j+1}$, where W^{t*}_{t+j} denotes self(t)'s perception about wealth in the future period $t+j$. Consider then the given selves' planned retirement savings s^{t*}_{P+1} and $s^{(t+1)*}_{P+1}$. By using Eq.(14) we get

$$s^{t*}_{P+1} = W^{t*}_P - \beta c^{t*}_t \overset{(c^{t*}_{t+1}<c_{t+1})}{>} W^{t*}_P - \beta c^{(t+1)*}_{t+1} \overset{(W^{t*}_P > W^{(t+1)*}_P)}{>} W^{(t+1)*}_P - \beta c^{(t+1)*}_{t+1} = s^{(t+1)*}_{P+1},$$

and hence for $i = NA$ and for all $t < P$ it holds that $s^{t*}_{P+1} > s^{(t+1)*}_{P+1}$. Thus, NA's planned retirement savings decreases in time and coincides with the actual retirement savings only in period P. From $C^{t*}_{SA} = C^t_{SA} \forall t \in \mathbb{N}^T$ it trivially follows for SA that $s^{t*}_{P+1} = s_{P+1}$ holds for all $t \in \mathbb{N}^P$.[52] For LN it

[52] Of course also $s^{t*}_{P+1+j} = s_{P+1+j}$ for any t and for any $j \in \mathbb{N}^{T-P-1}$.

is easy to verify $s_{P+1}^{t*} > s_{P+1}^{(t+1)*} > s_{P+1}$ since $c_{t+1}^{t*} < c_{t+1}^{(t+1)*}$ directly implies $W_P^{t*} > W_P^{(t+1)*}$.
iv) First, note that $\Delta s_{P+1}^{t*,NA} - \Delta s_{P+1}^{t*,LN} = s_{P+1}^{t*,NA} - s_{P+1}^{t*,LN}$ and rewrite then the claim as $s_{P+1}^{t*,NA} - s_{P+1}^{t*,LN} \geq 0 \; \forall t \leq P$. Pick an arbitrary self(t) for LN and NA. Write planned retirement savings in the following present value form

$$s_{P+1}^{t*,i} = Rs_t + \sum_{\tau=0}^{P-t}(w - c_{t+\tau}^{t*,i})\delta^\tau \text{ for } i \in \{NA, LN\}. \tag{20}$$

Elaborate Eq.(20) as

$$s_{P+1}^{t*,i} = Rs_t + w\sum_{\tau=0}^{P-t}\delta^\tau - \sum_{\tau=0}^{P-t} c_{t+\tau}^{t*,i}\delta^\tau,$$

and then reproduce the claim as

$$s_{P+1}^{t*,NA} - s_{P+1}^{t*,LN} = \sum_{\tau=0}^{P-t} c_{t+\tau}^{t*,LN}\delta^\tau - \sum_{\tau=0}^{P-t} c_{t+\tau}^{t*,NA}\delta^\tau \geq 0.$$

Then, write the LHS of the claim in the explicit form by using $c_{t+j}^{t*,NA} = c_t\beta$ and $c_{t+j}^{t*,LN} = \iota_{(t+j)}^{(t)} W_{t+j}$, and $\iota_t = (1 + \beta\sum_{\tau=1}^{T-t}\delta^\tau)^{-1}$ and $c_t = W_t\iota_t$. It results in

$$\begin{aligned} s_{P+1}^{t*,NA} - s_{P+1}^{t*,LN} &= c_t^{t*,LN} + \sum_{\tau=0}^{P-t} c_{t+\tau}^{t*,LN}\delta^\tau - c_t^{t*,NA}(1 + \beta\sum_{\tau=1}^{P-t}\delta^\tau) \\ &= W_t\iota_t + \sum_{\tau=0}^{P-t}\iota_{(t+j)}^{(t)} W_{t+j}\delta^\tau - \frac{W_t\iota_t}{\iota_{T-P+t}}. \end{aligned} \tag{21}$$

Since $\iota_t < \iota_{t+j}$ for all $j \in \mathbb{N}^{T-t}$, and hence $W_t\iota_t - \frac{W_t\iota_t}{\iota_{T-P+t}}$ is positive for all $t \in \mathbb{N}^T$, it follows that Eq.(21) is non-negative for all $\forall t \leq P$. ■

B.4 Proof of Proposition (3)

Proof. i) Assume $s_{k+1} > Rs_k$ for some $k < P$. The assumption holds only if $w - c_k > 0$, since $w - c_k + Rs_k \equiv s_{k+1}$, and thus $s_{k+1} > Rs_k \Leftrightarrow w - c_k > 0$. By Lemma (1) $c_t > c_{t+1}$ for all $t \in \mathbb{N}^{T-1}$, hence if $w - c_k > 0$ for some t then $w - c_{k+j} > 0$ for all $j \in \mathbb{N}^{P-k}$. It is then clear that once the agent starts positive saving, i.e. repayments or accumulation of the funds, he will never borrow again.

ii) Assume $t = \min\{k \leq P \,|\, s_{k+1} \geq Rs_k\}$ exists. Since now $k = t$ is the earliest period for which $s_{k+1} \geq Rs_k$, i.e. $w - c_k \geq 0$, it must be the case that $s_{n+1} < Rs_n$ for all $n \in \mathbb{N}^{t-1}$. Now, $s_{n+1} < Rs_n$ only if $w - c_n < 0$, and since $s_{n+1} < Rs_n$ for all $n \in \mathbb{N}^{t-1}$ it follows that $w - c_n < 0$ for all $n \in \mathbb{N}^{t-1}$. The agent does not have initial savings, hence we have $Rs_1 = 0$. Then, given that $s_{n+1} < Rs_n$ for all $n \in \mathbb{N}^{t-1}$ we have $Rs_n < 0$ for all $n \in \mathbb{N}_2^{t-1}$, and thus $Rs_n \leq 0$ for all $n \in \mathbb{N}^{t-1}$ which concludes the proof. ■

B.5 Proof of Lemma (2)

Proof. For an attractive SP for self(t) assume $k = p^{t*} \in \mathbb{N}_{t+1}^{P}$. Then $e > \left(\frac{W_t^{B_t}}{W_t^{B_k}}\right)^{\frac{\iota_t^{-1}}{1-\beta\delta^{k-t}}}$ and

$$k = \min\left\{\arg\max_{p\in\mathbb{N}_{t+1}^{P}} \left\{\hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right) + u\left(e\right)\left(1-\beta\delta^{p-t}\right)\middle| U_{t*}(C^{B_p}) - U_{t*}(C^{B_t}) > 0\right\}\right\}.$$

Now, gain from procrastination, i.e. $u\left(e\right)\left(1-\beta\delta^{p-t}\right)$, increases and is concave in p. Loss from procrastination is $\hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right)$ and can be written in the explicit form of $\iota_t^{-1}\ln\left(\frac{W_t^{B_p}}{W_t^{B_t}}\right)$, which is concave and decreasing in p. Hence, as a sum of two concave functions $\Delta U_t^p \equiv \hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right) + u\left(e\right)\left(1-\beta\delta^{p-t}\right)$ must be concave as well. So, it can be ever increasing, ever decreasing or first increasing and then decreasing in p. Since $p^{t*} \neq t$ ever decreasing is cancelled out as a probable form of ΔU_t^p here. If $p^{t*} = P$ it is clear that ΔU_t^p is ever increasing and condition $e > \left(\frac{W_t^{B_t}}{W_t^{B_n}}\right)^{\frac{\iota_t^{-1}}{1-\beta\delta^{n-t}}}$ holds for all $n \in \mathbb{N}_{t+1}^{k}$. If $p^{t*} \neq P$ it must be that p^{t*} is the last p for which ΔU_t^p is increasing since p^{t*}, by its definition, is the period for which ΔU_t^p is maximised. But then we must have ΔU_t^p is increasing for all $n < p$ as well, which completes the proof. ■

B.6 Proof of Lemma (3)

Proof. Pick any self(t), $t \in \mathbb{N}^{P-2}$. *Self(t) of ED* has intertemporal preferences of the form

$$U_t^{ED} = \sum_{\tau=0}^{T-t} \delta^{\tau} \ln\left(c_{\tau+t+1}\right) - \ln\left(e\right),$$

which implies plan $c_t = c_{t+\tau} \forall \tau \in \mathbb{N}^{T-t-1}$, and hence for any given W_t the optimal consumption is $c_{t+1}^{t*} = \varepsilon_{t+1} W_{t+1}$,where $\varepsilon_{t+1} = \frac{1}{\sum_{\tau=0}^{T-t-1}\delta^{\tau}}$. Now, $c_{t+1}^{t+1*} = \varepsilon_{t+1}W_{t+1} = \varepsilon_t W_t = c_{t+1}^{t*}$ since $c_{t+1}^{t+1*} = \varepsilon_{t+1}W_{t+1} = \varepsilon_{t+1}RW_t\left(1-\varepsilon_t\right) = \varepsilon_{t+1}RW_t\left(\frac{\sum_{\tau=0}^{T-t}\delta^{\tau}-1}{\sum_{\tau=0}^{T-t}\delta^{\tau}}\right) = \varepsilon_{t+1}RW_t\left(\frac{\sum_{\tau=1}^{T-t}\delta^{\tau}}{\sum_{\tau=0}^{T-t}\delta^{\tau}}\right) = \varepsilon_{t+1}W_t\left(\frac{\sum_{\tau=0}^{T-t-1}\delta^{\tau}}{\sum_{\tau=0}^{T-t}\delta^{\tau}}\right) = W_t\left(\sum_{\tau=0}^{T-t}\delta^{\tau}\right)^{-1} = W_t\varepsilon_t.$

Assume then that SP is attractive for self(t) but he wants to procrastinate for one period. We know that self(t) has temptation to procrastinates iff

$$U_t\left(C^{W_t^{B_t}}\right) = \sum_{\tau=0}^{T-t}\delta^{\tau}\ln\left(\varepsilon_t W_t^{B_t}\right) - \ln\left(e\right) < \sum_{\tau=0}^{T-t}\delta^{\tau}\ln\left(\varepsilon_t W_t^{B_{t+1}}\right) - \delta\ln\left(e\right) = U_t\left(C^{W_t^{B_{t+1}}}\right)$$

$$\Rightarrow \sum_{\tau=0}^{T-t}\delta^{\tau}\ln\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right) < \left(1-\delta\right)\ln\left(e\right).$$

Self(t+1) will then start SP if it is not profitable to procrastinate it, i.e. iff

$$
\begin{aligned}
U_{t+1}\left(C^{W_{t+1}^{B_{t+1}}}\right) &= \sum_{\tau=0}^{T-t-1}\delta^{\tau}\ln\left(\varepsilon_{t+1}W_{t+1}^{B_{t+1}}\right)-\ln(e) \geq \sum_{\tau=0}^{T-t-1}\delta^{\tau}\ln\left(\varepsilon_{t+1}W_{t+1}^{B_{t+2}}\right)-\delta\ln(e)\\
&= U_{t+1}^{t}\left(C^{W_{t+1}^{B_{t+2}}}\right)\\
&\Rightarrow \sum_{\tau=0}^{T-t-1}\delta^{\tau}\ln\left(\frac{W_{t+1}^{B_{t+1}}}{W_{t+1}^{B_{t+2}}}\right) \geq (1-\delta)\ln(e).
\end{aligned}
$$

But then this implies $\sum_{\tau=0}^{T-t-1}\delta^{\tau}\ln\left(\frac{W_{t+1}^{B_{t+1}}}{W_{t+1}^{B_{t+2}}}\right) > \sum_{\tau=0}^{T-t}\delta^{\tau}\ln\left(\frac{W_{t}^{B_{t}}}{W_{t}^{B_{t+1}}}\right)$ and to get a contradiction it is enough to show $\ln\left(\frac{W_{t+1}^{B_{t+1}}}{W_{t+1}^{B_{t+2}}}\right) \leq \ln\left(\frac{W_{t}^{B_{t}}}{W_{t}^{B_{t+1}}}\right)$, i.e.that $\ln\left(\frac{W_{t}^{B_{t}}}{W_{t}^{B_{t+1}}}\right)$ is decreasing in t. Let us assume that $W_t > 1\ \forall t \in \mathbb{N}^P$. Then analogously with the proof of Lemma (**??**) it is enough to show that $W_t^{B_t} - W_t^{B_{t+1}}$ is decreasing in t. Now,

$$
\begin{aligned}
W_t^{B_t} - W_t^{B_{t+1}} &= Rs_t + \sum_{\tau=0}^{P-t}\delta^{\tau}(w-\gamma) + \delta^{P+1-t}\sum_{\tau=1}^{P+1-t}R_B^{\tau}\gamma\\
&\quad -\left(Rs_t + \sum_{\tau=0}^{P-t}\delta^{\tau}(w-\gamma) + \gamma + \delta^{P+1-t}\sum_{\tau=1}^{P-t}R_B^{\tau}\gamma\right)\\
&= \gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P+1-t}R_B^{\tau} - \sum_{\tau=0}^{P-t}\delta^{\tau}\right) - \gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P-t}R_B^{\tau} - \sum_{\tau=1}^{P-t}\delta^{\tau}\right)\\
&= \gamma\left(\delta^{P+1-t}R_B^{P+1-t}-1\right) = \gamma\left((\delta R_B)^{P+1-t}-1\right).
\end{aligned}
$$

Clearly, $(\delta R_B)^{P+1-t} - 1$ is decreasing in t since $P+1-t$ is decreasing in t. Hence, $\ln\left(\frac{W_{t}^{B_{t}}}{W_{t}^{B_{t+1}}}\right)$ is decreasing in t and we have a contradiction, which completes the proof. ∎

B.7 Proof of Lemma (4)

Proof. Let us show that such utility cost $u(e)$ exists for which Eq.(**??**) holds simultaneously with Eq.(**??**). By noting $1-\beta\delta > 1-\delta$ we find that if $\hat{U}_{(t+1)*}(C^{B_{t+1}}) - \hat{U}_{(t+1)*}(C^{B_{t+2}}) \leq \hat{U}_{t+1}^{t*}(C^{B_{t+1}}) - \hat{U}_{t+1}^{t*}(C^{B_{t+2}})$ holds, then that is a sufficient condition to prove our claim. Starting from the right hand side we can elaborate its parts as follows

$$
\begin{aligned}
\hat{U}^{t*}_{t+1}(C^{B_{t+1}}) &= \sum_{\tau=0}^{T-t-1} \delta^\tau \ln(\iota_t W^{B_{t+1}}_{t+1}\beta) = \ln(\iota_t W^{B_{t+1}}_{t+1}\beta) \sum_{\tau=0}^{T-t-1} \delta^\tau, \\
\hat{U}^{t*}_{t+1}(C^{B_{t+2}}) &= \sum_{\tau=0}^{T-t-1} \delta^\tau \ln(\iota_t W^{B_{t+2}}_{t+1}\beta) = \ln(\iota_t W^{B_{t+2}}_{t+1}\beta) \sum_{\tau=0}^{T-t-1} \delta^\tau, \\
\Rightarrow \quad \hat{U}^{t*}_{t+1}(C^{B_{t+1}}) - \hat{U}^{t*}_{t+1}(C^{B_{t+2}}) &= \ln\left(\frac{W^{B_{t+1}}_{t+1}}{W^{B_{t+2}}_{t+1}}\right) \sum_{\tau=0}^{T-t-1} \delta^\tau. \qquad (22)
\end{aligned}
$$

Similar elaboration for the left hand side produces

$$
\begin{aligned}
\hat{U}_{(t+1)*}(C^{B_{t+1}}) &= \ln(\iota_{t+1} W^{B_{t+1}}_{t+1}) + \beta \sum_{\tau=1}^{T-t-1} \delta^\tau \ln(\iota_{t+1} W^{B_{t+1}}_{t+1}\beta), \\
\hat{U}_{(t+1)*}(C^{B_{t+2}}) &= \ln(\iota_{t+1} W^{B_{t+2}}_{t+1}) + \beta \sum_{\tau=1}^{T-t-1} \delta^\tau \ln(\iota_{t+1} W^{B_{t+2}}_{t+1}\beta), \\
\Rightarrow \quad \hat{U}_{(t+1)*}(C^{B_{t+1}}) - \hat{U}_{(t+1)*}(C^{B_{t+2}}) &= \ln\left(\frac{W^{B_{t+1}}_{t+1}}{W^{B_{t+2}}_{t+1}}\right) + \beta \ln\left(\frac{W^{B_{t+1}}_{t+1}}{W^{B_{t+2}}_{t+1}}\right) \sum_{\tau=1}^{T-t-1} \delta^\tau. \qquad (23)
\end{aligned}
$$

Then, subtracting the R.H.S. of Eq (22) from the R.H.S. of Eq (23) results in

$$
\left((1-\beta) \sum_{\tau=1}^{T-t-1} \delta^\tau\right) \ln\left(\frac{W^{B_{t+1}}_{t+1}}{W^{B_{t+2}}_{t+1}}\right) > 0,
$$

since $1-\beta > 0$ and $W_t^{B_t} > W_t^{B_{t+i}}$ $\forall t, i \in \{1, 2, ..., P-t\}$. ■

B.8 Proof of Lemma (5)

Proof. Assume $W_t \geq e$ $\forall t \in \mathbb{N}^P$ and the revised set up. Pick an arbitrary self(t) such that $t \in \mathbb{N}^P$. Then, SP B_t is attractive for self(t) iff

$$
\begin{aligned}
U_{t*}(C) \quad &< \quad U_{t*}(C^{B_t}) = \hat{U}_{t*}(C^{B_t}) - u(e) \\
&\Leftrightarrow \quad \ln(\iota_t W_t) + \beta \sum_{\tau=1}^{T-t} \delta^\tau \ln(\beta \iota_t W_t) > \ln(\iota_t W_t^{B_t}) + \beta \sum_{\tau=1}^{T-t} \delta^t \ln(\beta \iota_t W_t^{B_t}) - \ln(e) \\
&\Leftrightarrow \quad \ln(e) < \left(1 + \beta \sum_{\tau=1}^{T-t} \delta^\tau\right) \ln\left(\frac{W_t^{B_t}}{W_t}\right) = \iota_t^{-1} \ln\left(\frac{W_t^{B_t}}{W_t}\right) \\
&\Leftrightarrow \quad e < \left(\frac{W_t^{B_t}}{W_t}\right)^{\iota_t^{-1}}
\end{aligned}
$$

Now, it is enough to show that $\iota_t^{-1} \ln\left(\frac{W_t^{B_t}}{W_t}\right)$ is decreasing in t. Clearly ι_t^{-1} is decreasing in t, and hence if then also $\ln\left(\frac{W_t^{B_t}}{W_t}\right)$ is decreasing in t we are done. To show that $\ln\left(\frac{W_t^{B_t}}{W_t}\right)$ is decreasing in t we have to show that $\frac{W_t^{B_t}}{W_t}$ is decreasing in t. Since $W_t \geq e$ and thus $\ln(W_t) \geq \ln(e)$, we clearly have $W_t \geq 1$ due to construction of the revised set up and especially by the assumption of $u(e) > 0$. Hence, we know that $W_t^{B_t} > W_t \geq 1$ $\forall t \in \mathbb{N}^P$, from which it then follows that now $\forall t \in \mathbb{N}^P$ holds $\ln(W_t^{B_t}) - \ln(W_t) \leq W^{B_t} - W_t$ by the concavity of the logarithmic function. Hence, if $W_t^{B_t} - W_t$ is decreasing in t then necessarily $\ln(W_t^{B_t}) - \ln(W_t) = \ln \frac{W_t^{B_t}}{W_t}$ is decreasing in t, which in its part implies that $\frac{W_t^{B_t}}{W_t}$ is decreasing in t. Now,

$$
\begin{aligned}
\Delta_t &\equiv W_t^{B_t} - W_t = Rs_t + \sum_{\tau=0}^{P-t} \delta^\tau (w-\gamma) + \delta^{P+1-t} \sum_{\tau=1}^{P+1-t} R_B^\tau \gamma - \left(Rs_t + \sum_{\tau=0}^{P-t} \delta^\tau w\right) \\
&= \gamma \left(\delta^{P+1-t} \sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t} \delta^\tau\right) \\
&\Rightarrow \Delta_{t+1} = \gamma \left(\delta^{P-t} \sum_{\tau=1}^{P-t} R_B^\tau - \sum_{\tau=0}^{P-t-1} \delta^\tau\right).
\end{aligned}
$$

Hence, we get

$$
\begin{aligned}
\Delta_t - \Delta_{t+1} &= \gamma \left(\delta^{P+1-t} \sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t} \delta^\tau\right) - \gamma \left(\delta^{P-t} \sum_{\tau=1}^{P-t} R_B^\tau - \sum_{\tau=0}^{P-t-1} \delta^\tau\right) \\
&= \gamma \left(\delta^{P+1-t} \sum_{\tau=1}^{P+1-t} R_B^\tau - \delta^{P-t} \sum_{\tau=1}^{P-t} R_B^\tau - \delta^{P-t}\right) \\
&= \gamma \left(\delta^{P-t} \left(\delta \sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=1}^{P-t} R_B^\tau\right) - \delta^{P-t}\right) \\
&= \gamma \delta^{P-t} \left(\delta \sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=1}^{P-t} R_B^\tau - 1\right) \\
&\Rightarrow \Delta_t - \Delta_{t+1} \geq 0, \text{ i.e. } W_t^B - W_t \text{ is decreasing,} \\
&\Leftrightarrow \delta \sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=1}^{P-t} R_B^\tau - 1 \geq 0 \\
&\Leftrightarrow \delta \geq \frac{1 + \sum_{\tau=1}^{P-t} R_B^\tau}{\sum_{\tau=1}^{P+1-t} R_B^\tau} = \frac{1 + \sum_{\tau=1}^{P+1-t} R_B^\tau - R_B^{P+1-t}}{\sum_{\tau=1}^{P+1-t} R_B^\tau} = 1 + \frac{1 - R_B^{P+1-t}}{\sum_{\tau=1}^{P+1-t} R_B^\tau} \\
&= 1 + \frac{1 - R_B^{P+1-t}}{R_B \frac{1 - R_B^{P+1-t}}{1 - R_B}} = 1 + \frac{1 - R_B}{R_B} = \frac{1}{R_B},
\end{aligned}
$$

which holds given the revised set up since by definition $r < r_B$, and hence $\delta \equiv \frac{1}{1+r} > \frac{1}{1+r_B} \equiv \frac{1}{R_B}$ for all given δs and R_Bs. ■

B.9 Proof of Proposition (4)

Proof. Assume that SP has not been started, and that it is still attractive for self(t), i.e.

$$\left(\frac{W^{B_t}}{W_t}\right)^{\iota_t^{-1}} > e.$$

By Lemma (**??**) we know that

$$\left(\frac{W^{B_1}}{W_1}\right)^{\iota_1^{-1}} > \left(\frac{W^{B_t}}{W_t}\right)^{\iota_t^{-1}} > e,$$

i.e. SP is certainly attractive for self(1) as well.

The optimal starting period is given by

$$p^{t*} \equiv \min\left\{ \arg\max_{p\in\mathbb{N}_{t+1}^{P}} \left\{ \hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right) + u\left(e\right)\left(1-\beta\delta^{p-t}\right) \middle| U_{t*}(C^{B_p}) - U_{t*}(C^{B_t}) \geq 0 \right\} \right\}$$

where $\hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right)$ can be interpreted as total loss and $u\left(e\right)\left(1-\beta\delta^{p-t}\right)$ as total gain from procrastination. Procrastination is worthwhile from self(t)*'s* viewpoint if

$$\hat{U}_{t*}\left(C^{B_p}\right) - \hat{U}_{t*}\left(C^{B_t}\right) + u\left(e\right)\left(1-\beta\delta^{p-t}\right) = \iota_t^{-1}\ln\left(\frac{W_t^{B_p}}{W_t^{B_t}}\right) + \left(1-\beta\delta^{p-t}\right)\ln\left(e\right) > 0.$$

By Lemma (2) we know that procrastination is tempting if it is tempting for one period, i.e. if $\frac{\iota_t^{-1}}{(1-\beta\delta)}\ln\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right) < \ln\left(e\right)$. Hence, procrastination is tempting for self(t) if

$$\begin{aligned} &\iota_t^{-1}\ln\left(\frac{W_t^{B_{t+1}}}{W_t^{B_t}}\right) + \left(1-\beta\delta\right)\ln\left(e\right) > 0 \\ \Leftrightarrow\quad &\iota_t^{-1}\ln\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right) < \left(1-\beta\delta\right)\ln\left(e\right). \end{aligned} \tag{24}$$

Now, if $W_t \geq e > 1$ for all $t \in \mathbb{N}^P$

$$\ln\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right) = \ln\left(W_t^{B_t}\right) - \ln\left(W_t^{B_{t+1}}\right) < \left(W_t^{B_t} - W_t^{B_{t+1}}\right) = \gamma\left(\left(\delta R_B\right)^{P+1-t} - 1\right).$$

Clearly, RHS of the inequality is decreasing in t, which implies LHS of it is decreasing in t. Since ι_t^{-1} is decreasing in t, $\iota_t^{-1}\ln\left(\frac{W_t^{B_t}}{W_t^{B_{t+1}}}\right)$ is certainly decreasing in t, and hence it follows that if condition (24) holds for self(1) it holds for every subsequent self too. This, from its part, means that if self(1) does not start SP there's no self who would start SP while at least one period is procrastinated since self(1) procrastinates his starting of SP. ■

B.10 Proof of Lemma (6)

Proof. We have to show that $\mathcal{A}_1^{SA} = \mathcal{A}_1^{NA}$. To show this we first show that $W_{t+1}^{B_t} = R\left(W_t^{B_t} - \iota_t W_t^{B_t}\right)$. If the savings program is started in the period t the wealth in the next period will clearly be

$$W_{t+1}^{B_t} = R\left(W_t - \gamma - \iota_t W_t^{B_t}\right) + \gamma\left(\delta^{P-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t-1}\delta^\tau\right), \tag{25}$$

where

$$\begin{aligned} W_t &= Rs_t + w\sum_{\tau=0}^{P-t}\delta^\tau, \text{ and} \\ W_t^{B_t} &= Rs_t + (w-\gamma)\sum_{\tau=0}^{P-t} + \delta^{P+1-t}\gamma\sum_{\tau=1}^{P+1-t} R_B^\tau = W_t + \gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t}\delta^\tau\right) \end{aligned}$$

by the construction of the revised set up. Simple elaboration of Eq.(25) results in

$$\begin{aligned} W_{t+1}^{B_t} &= R\left(W_t - \gamma - \iota_t W_t^{B_t}\right) + \gamma\left(\delta^{P-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t-1}\delta^\tau\right) \\ &= R\left(W_t - \gamma - \iota_t W_t^{B_t} + \gamma\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \delta\gamma\sum_{\tau=0}^{P-t-1}\delta^\tau\right) \\ &= R\left(W_t - \iota_t W_t^{B_t} + \gamma\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \gamma\sum_{\tau=0}^{P-t}\delta^\tau\right) \\ &= R\left(\underbrace{W_t + \gamma\left(\delta^{P+1-t}\sum_{\tau=1}^{P+1-t} R_B^\tau - \sum_{\tau=0}^{P-t}\delta^\tau\right)}_{W_t^{B_t}} - \iota_t W_t^{B_t}\right) \\ &= R\left(W_t^{B_t} - \iota_t W_t^{B_t}\right) \\ &\Rightarrow\ W_{t+1}^{B_t} = R\left(W_t^{B_t} - \iota_t W_t^{B_t}\right). \end{aligned} \tag{26}$$

Now,

$$\mathcal{A}_1^{SA} = \frac{W_1^{B_1}}{W_1}\prod_{\tau=1}^{T-t}\left(\frac{W_{\tau+1}^{B_1}}{W_{\tau+1}}\right)^{\beta\delta^\tau}. \tag{27}$$

By using Eq.(26) we get

$$\frac{W_2^{B_1}}{W_2} = \frac{R\left(W_1^{B_1} - \iota_1 W_1^{B_1}\right)}{R\left(W_1 - \iota_1 W_1\right)} = \frac{RW_1^{B_1}\left(1-\iota_1\right)}{RW_1\left(1-\iota_1\right)} = \frac{W_1^{B_1}}{W_1} \tag{28}$$

Reapplying Eq.(26) with Eq.(28) we get

$$\frac{W_1^{B_1}}{W_1} = \frac{W_{1+\tau}^{B_1}}{W_{1+\tau}} \text{ for all } \tau \in \mathbb{N}^{T-1}.$$

Plugging this solution in to the attractiveness condition (27) yields

$$\begin{aligned} \mathcal{A}_1^{SA} &= \frac{W_1^{B_1}}{W_1} \prod_{\tau=1}^{T-t} \left(\frac{W_1^{B_1}}{W_1} \right)^{\beta\delta^\tau} = \left(\frac{W_1^{B_1}}{W_1} \right)^{1+\beta \sum_{\tau=1}^{T-t} \delta^\tau} \\ &= \left(\frac{W_1^{B_1}}{W_1} \right)^{\iota_t^{-1}} = \mathcal{A}_1^{NA} \\ &\Rightarrow \mathcal{A}_1^{SA} = \mathcal{A}_1^{NA}, \end{aligned}$$

which completes the proof. ■

B.11 Proof of Lemma (7)

Proof. Assume that SP is attractive for self(t) of SA. If $t = P$ it is trivial that self(t) starts the attractive SP immediately. Hence, suppose that $t < P$. Now, self(t) prefers starting of the attractive SP immediately to never starting it. If he uses the manipulative consumption in the case where it results in $\mathcal{A}_{t+1}^{SA} \leq e$, self($t$) knows this and hence prefers not to use the manipulative consumption, since using it would cause self*(t+1)* to abandon SP for good. If self(t) prefers not to use the manipulative consumption it must be the case that he does not procrastinate but starts SP immediately. ■

B.12 Proof of Proposition (5)

Proof. i) Assume that there exists self*(k)*, $k \in \mathbb{N}_{t+1}^{p_k^{tol}}$, who starts the savings program if self(t) procrastinates. Assume then that self(t) prefers immediate starting of SP to delaying starting of SP, i.e. procrastination is intolerable for self(t). Then it must be the case that $U\left(C_t^{mc}\right) - U_{t*}\left(C^{W^{B_t}}\right) < 0$ for all $p_t \in \mathbb{N}_{t+1}^{P}$, but then there can not exist self*(k)*, $k \in \mathbb{N}_{t+1}^{p_t^{tol}}$. Contradiction.
ii) Point (i) and Lemma (7). ■

B.13 Proof of Corollary (2)

Proof. i) Assume the revised set up, $i = SA$, let $e < \mathcal{A}_t^{SA}$, and suppose that SP is non-started. By the analogy between NA's and SA's choices when the attractive SP is not started we know that if $e < \mathcal{A}_t^{SA}$ then $e < \mathcal{A}_1^{SA}$. If SP is attractive for self(1) of SA then by Proposition (5) he either starts it immediately or then if he procrastinates it will be started in period $k \in \mathbb{N}_2^{p_1^{tol}}$.
ii) Suppose not,i.e. suppose that the retirement savings are greater in the absence than in the presence of SP. Then, from self(1)'s perspective the discounted utility from the retirement time, $\tau \in \mathbb{N}_{P+1}^{T}$, must be higher in the absence than in the presence of SP. Now, for any wealth level posted for the retirement, W_{P+1}^{Ret} the optimal consumption during the retirement time is given

by the fixed sequence of MPCs multiplied by the relevant periodical wealth, i.e consumption is given by $\left\{\iota_\tau W_\tau^{\mathrm{Ret}}\right\}_{\tau=P+1}^{T}$. This sequence produces higher utility only if $W_{P+1}^{\mathrm{Ret}} > W_{P+1}^{\mathrm{Ret},B_{p^*}}$. But then, use of fixed MPCs also for working periods, $\tau \in \mathbb{N}^P$, results in sequences $\left\{\iota_\tau W_\tau^{\mathrm{Ret}}\right\}_{\tau=1}^{P}$ and $\left\{\iota_\tau W_\tau^{\mathrm{Ret},B_{p^*}}\right\}_{\tau=1}^{P}$, where $\iota_\tau W_\tau^{\mathrm{Ret}} > \iota_\tau W_\tau^{\mathrm{Ret},B_{p^*}}$ for all $\tau \in \mathbb{N}^P$. Particularly, now $W_1^{\mathrm{Ret}} > W_1^{\mathrm{Ret},B_{p^*}}$, which is a contradiction, since by the claim (i) the savings program is attractive for self(1). Recall, $\ln(e) > 0 \Rightarrow e > 1$, and SP is attractive for self(1) iff $\mathcal{A}_1 > e$ which is equivalent with $\ln\left(\frac{W_1^{B_{p^*}}}{W_1}\right)^{\iota_1^{-1}} > \ln(e)$. Thus, for the attractiveness for self(1) must hold $W_1^{B_{p^*}} > W_1$ since $\iota_\tau^{-1} > 1$ for all $\tau \in \mathbb{N}^{T-1}$. ■

B.14 Proof of Proposition (6)

Proof. (i)-(ii) Analogy with the behaviour of *NA*'s self(1), Proposition (4), and Eqs.(11) and (12). (iii) Let $\mathcal{A}_t > e$ and $t \geq 3$. Then, given self(t)'s perception about future behaviour, he starts SP iff there does not exist self(k), $k \in \mathbb{N}_{t+1}^{p_t^{tol}}$, and otherwise procrastinates. For the first case, suppose that defined self(k) does not exists. Then, after procrastination of $t-1$ periods, self(t) starts SP and hence by point (ii) in Corollary (2) retirement savings are greater in the presence than in the absence of SP. For the other case, suppose that defined self(k), $k \in \mathbb{N}_t^{p_{t-1}^{tol}}$, does exists for self($t-1$), and note that the existence is based on self($t-1$)'s perception about future behaviour. In addition and with out loss of generality, suppose that $\iota_{t-1} W_{t-1}^{B_t} \Rightarrow \mathcal{A}_t^{LN} \leq e$, i.e. suppose that self($t-1$) is the last self for whom SP is attractive and hence use of the manipulative consumption leads abandoning of SP in the subsequent period. Now, self($t-1$) procrastinates and self(t) abandons SP for good; thus, SP has been procrastinated $t-1$ periods and due to use of manipulative consumption retirement savings will be lower in the presence than in the absence of SP. ■