DICE

DISCUSSION PAPER

NO 384

Planning and Saving for Retirement

Tomasz Sulka

March 2022

IMPRINT

DICE DISCUSSION PAPER

Published by:
Heinrich-Heine-University Düsseldorf,
Düsseldorf Institute for Competition Economics (DICE),
Universitätsstraße 1, 40225 Düsseldorf, Germany
www.dice.hhu.de

Editor:
Prof. Dr. Hans-Theo Normann
Düsseldorf Institute for Competition Economics (DICE)
Tel +49 (0) 211-81-15125, E-Mail normann@dice.hhu.de

ISSN 2190-9938 (online) / ISBN 978-3-86304-383-4

Planning and Saving for Retirement*

Tomasz Sulka†

March 2022

Abstract

Planning for retirement and subsequent execution of the plan are difficult, but essential for financial security in old age. To formally analyse the interplay between planning and self-control, I introduce cognitive costs of formulating a plan into the dual-self model of impulse control. The resulting model can generate rational inaction in pension choices, with the agent's self-control and level of income playing a role of inputs into the decision whether or not to undertake costly planning. Furthermore, when they do plan, agents characterised by poor self-control save over shorter horizons and accumulate lower pension wealth. The possibility of rational inaction can explain other robustly observed behaviours, such as disproportionately low savings of individuals on low incomes and non-fungibility between public and private pension wealth. The model is applied to study welfare and savings implications of automatic enrolment into private pensions. The default option effect on plan participation arises due to the fact that counterfactual non-savers have the lowest threshold for accepting the default scheme. Nevertheless, the impact of automatic enrolment on total savings is ambiguous in general, because in addition to the counterfactual non-savers, the default may anchor contributions of a counterfactual active saver to a low default contribution rate. Consequently, although raising the default contribution rate itself has an ambiguous impact on aggregate savings, it always reduces the dispersion in pension wealth accumulation.

JEL: D14, D15, D91, E21, E71, H55, J32.
Keywords: Planning; Self-Control; Cognitive Costs; Pensions; Automatic Enrolment

*I am indebted to my thesis advisors, Philipp Kircher and Jan Grobovšek, for their continuous support, advice, and kindness. I would like to thank Paul Heidhues, Ed Hopkins, Bernhard Kasberger, Mats Köster, Botond Kőszegi, and Marco Schwarz for helpful discussions and audiences at the Annual PhD Conference in Leicester, the European Economic Association Congress in Geneva and in Rotterdam, and the University of Edinburgh for their comments. Financial support of the ESRC PhD Studentship is gratefully acknowledged.

†DICE, University of Düsseldorf. Email: sulka@dice.hhu.de.

1 Introduction

Financial preparation for retirement appears to be a difficult task for an average individual. This has recently become an important policy concern, as in most developed countries the ongoing reforms tend to reduce the generosity of public pension benefits and increase the importance of private pension arrangements. Because the latter are predominantly of the defined-contribution type, this shift endows individuals with much greater responsibility for their financial security in retirement (OECD, 2021). The challenges underlying retirement preparation can be separated into two distinct categories. First, an individual must come up with an appropriate plan that specifies, broadly speaking, a suitable savings vehicle and a desired savings rate. Second, an individual should exert enough financial self-control in order to implement the plan over time. Both requirements are potentially problematic and thus failures to save adequately seem widespread (Thaler and Benartzi, 2004; Benartzi and Thaler, 2007). Importantly, these two stages of financial preparation for retirement are conceptually distinct in the sense that they impose different requirements on a saver. For example, while the task of 'planning' requires sufficient financial knowledge and ability to process large amounts of information, the task of 'executing the plan' requires continuous attention to one's finances and ability to resist temptation to spend. However, the existing literature does not provide a formal framework for analysing the interplay between costly planning and exercising self-control to follow through with the plan. In this paper, I therefore introduce the cognitive costs of planning into a model of costly self-control.

To capture myopic deviations from a devised plan of action, I adopt a two-system, or dual-self, model of impulse control (Benhabib and Bisin, 2005; Fudenberg and Levine, 2006, 2012; Brocas and Carrillo, 2008) and modify it in two ways. First, I propose a frame under which 'savings targets' represent a financial plan and play a role of an internal commitment device. Using financial goals in order to control the temptation to spend is a central idea behind the theory of mental accounting (Thaler, 1985, 1999; Shefrin and Thaler, 1988). Second, in order to account for the mental effort of planning for retirement, I introduce the 'cognitive costs' of formulating a plan. These costs are meant to represent disutility from mental and psychological strain associated with acquiring relevant knowledge and processing large amounts of information, both of which are necessary for financial planning. The idea to model difficult decisions as cognitively costly is not new or particularly controversial (Conlisk, 1996; Reis, 2006; Alaoui and Penta, 2016; Choukhmane, 2021), and it is motivated by the empirical findings regarding the propensity to plan (Ameriks, Caplin, and Leahy, 2003; Lusardi and Mitchell, 2007, 2011; Van Rooij, Lusardi, and Alessie, 2012), the importance of financial literacy (Lusardi and Mitchell, 2014; Lusardi, Michaud, and Mitchell, 2017), and the impact of cognitive ability on economic decision-making (Banks and Oldfield, 2007; Banks, o'Dea, and Oldfield, 2010; Smith, McArdle, and Willis, 2010).

In the model, an agent faces a finite-horizon problem of choosing life-cycle savings. The decision when, if at all, to bear the cognitive cost of planning for retirement is based on the assessment of lifetime utility gains from consumption smoothing. As a direct consequence, costly planning allows for the emergence of *rational inaction* in pension choices, which occurs when the cognitive cost exceeds potential utility gains from consumption smoothing. Because these gains are partly determined by the agent's ability to follow through with the plan (i.e., self-control), the model implies that non-savers are characterised by poor financial self-control and high enough cognitive costs. Furthermore, when they do plan, agents characterised by poor self-control save over shorter horizons and accumulate lower pension wealth.

The possibility of rational inaction can also explain other behaviours robustly observed in the domain of retirement savings, such as disproportionately low savings of individuals on low income (Dynan, Skinner, and Zeldes, 2004). Moreover, the presence of cognitive costs of planning generates a particular form of non-fungibility between public and private pension wealth. Namely, an exogenous increase in a public pension benefit can crowd out private savings more than proportionately, if a reduced incentive to save results in the agent planning over a shorter horizon, or not at all.

Regarding the theory's policy implications, I focus on the effects of automatic enrolment into private pensions, an increasingly popular policy tool designed to take advantage of widespread inertia in pension choices. The model rationalises a large positive impact of automatic enrolment on plan participation (the 'default option effect'). That is due to the fact that counterfactual non-savers have the lowest threshold for perceiving automatic enrolment as welfare-improving and thus do not opt out, even if the underlying pension scheme is 'imperfect' in the sense that it imposes an implicit cost on the saver (e.g., by being actuarially unfair or not well calibrated to the agent's first-best saving schedule). The impact of automatic enrolment on total savings is ambiguous, however. That is because in addition to the counterfactual non-savers, the default may anchor savings of a counterfactual active saver to a low default contribution rate. Consequently, although raising the default contribution rate itself has an ambiguous impact on aggregate savings, it always reduces the dispersion in pension wealth accumulation. All three effects have been robustly observed in the field (Madrian and Shea, 2001; Thaler and Benartzi, 2004; Choi, Laibson, Madrian, Metrick, and Poterba, 2004) and the theoretical framework proposed here provides a lens through which this empirical evidence may be interpreted.

If accounting for inaction is essential for realistic predictions in the domain of retirement savings, why not refer to a well-known model of procrastination by O'Donoghue and Rabin (1999a, 2001)? In essence, the model of O'Donoghue and Rabin produces procrastination due to the agent's naiveté about the self-control problem. At time t, the agent postpones an unpleasant (costly) activity, because he overestimates his willingness to carry out the task in the following period. However, when period $t+1$ comes, the activity seems just as unpleasant and the agent postpones again. This mechanism is fundamentally different from the model of rational inaction proposed here. There is an important distinction to be made between procrastinating on a known best action (as in O'Donoghue and Rabin) and inaction due to the inability to infer the optimal choice (as in the present paper). Arguably, the latter mechanism is better suited to modelling retirement-related decisions, while the model based on naiveté about self-control is more applicable to situations such as (not) going to the gym or unhealthy eating. Even though the two models can generate similar patterns of behaviour, taking a stand on the underlying mechanism is crucial for the model's comparative statics and policy implications. I draw a more precise comparison between the two theories in section 2.3 and discuss some relevant evidence throughout the paper.

Related literature and contribution. This paper contributes to three strands of the literature. First, there is convincing empirical evidence that non-planning and thus lack of active decision-making in preparation for retirement are widespread (Ameriks et al., 2003; Lusardi and Mitchell, 2007, 2011; Van Rooij et al., 2012). This paper contributes to the theoretical literature on personal plans (Beshears, Milkman, and Schwartzstein, 2016; Galperti, 2019) by identifying some novel determinants of non-planning, namely the agent's self-control and her level of income. Moreover, the multiperiod model analysed here sheds some light on the dynamics of planning and saving over a finite horizon.

Second, without a formal model it is difficult to interpret the available empirical evidence and assess the impact of automatic enrolment into pensions on welfare and other components of the household's balance sheet (Beshears, Choi, Laibson, and Madrian, 2018a; Bernheim and Taubinsky, 2018). By proposing a concrete theory of inaction in pension choices, this paper offers a lens through which this evidence may be interpreted. Specifically, under rational inaction, a default is only sticky when it constitutes a welfare improvement, although agents with high cognitive costs of planning and/or poor self-control are prone to accepting defaults that are suboptimal for them. Furthermore, the model yields precise predictions regarding the extent of crowding out of private savings by automatically collected pension contributions.[1]

Third, motivated by the intuitive ideas behind the theory of mental accounting (Thaler, 1985, 1999; Shefrin and Thaler, 1988), I model self-control as exerted via an internal commitment device in a form of savings targets. Capturing those ideas within a theoretical framework of a two-system model (Benhabib and Bisin, 2005; Fudenberg and Levine, 2006, 2012; Brocas and Carrillo, 2008) allows me to bridge a gap between those abstract models of impulse-control

[1] Choi, Laibson, Madrian, and Metrick (2003) derive a socially optimal default contribution rate when the agents are naïve procrastinators and have heterogeneous optimal savings rates. Goldin and Reck (2020) point out that depending on the underlying model of decision-making, the barriers to selecting an option that is not the default may or may not be treated as welfare-relevant by the social planner. They characterise the optimal default conditional on normativity of the decision-making costs and the propensity of an individual to make a mistake when choosing actively.

and the original work on mental accounting, which introduced the category-specific financial goals and non-fungibility of various forms of wealth by assumption and did not account for non-planning.[2]

The remainder of this paper is structured as follows. Section 2 outlines the model and analyses the implications of rational inaction for savings behaviour. Section 3 focuses on automatic enrolment into private pensions. Section 4 concludes. The proofs are relegated to the supplementary appendix.

2 The model

2.1 Setup

Economic environment. Time is discrete and finite, with T periods indexed by $t = 1, 2, ..., T$. The agent is employed in the first $T - 1$ periods, earning disposable income of y_t in period t. The last period of the model illustrates retirement. The timing of retirement and the level of income are exogenous.

During the working life, the agent can save into an illiquid account with a gross rate of return $R \geq 1$ per period, which can only be accessed in period T. No borrowing against this pension wealth is permitted. Denoting the agent's savings into the illiquid account in period t by a_t, the pension wealth available to them in retirement is:

$$A_T = y_T + \sum_{t=1}^{T-1} R^{T-t} a_t$$

where $a_t \geq 0$ for $t = 1, 2, ..., T - 1$ and $y_T \geq 0$ is a baseline level of (public) pension income provided exogenously to an agent, independently of their private savings. To focus on the life-cycle aspect of savings choices, the model is deterministic, i.e. there is no uncertainty regarding the agent's survival, working-life income, or returns to pension wealth.

Preferences. The goal of this paper is to develop a model of decision-making where 'planning' is followed by 'execution of the plan'. To account for a potential self-control problem at the plan execution stage, I adopt a two-system framework, which explicitly accounts for a conflict between the agent's myopic and forward-looking objectives. In each period, a 'doer' (the economic agent's impatient, myopic self, "he") would like to maximise the instantaneous utility, while a 'planner' (the economic agent's far-sighted, rational self, "she") would like to maximise the lifetime utility.

[2]In a recent contribution, Kőszegi and Matějka (2020) develop a model of mental accounting based on costly attention, which generates non-fungibility across spending categories, rather than imperfect substitutability between different sources of wealth yielding identical ex post utility.

As in Fudenberg and Levine (2006), I focus on the outcomes of a game played between the (long-lived) planner and a sequence of completely myopic doers.

Within each period, the doer ultimately controls the disposable income, but the planner can exert her influence by exercising costly self-control. Bridging the literatures on personal plans and costly self-regulation, I propose that savings targets, i.e. financial goals set by an individual for herself, play a role of an internal commitment device that allows the planner to discipline doer's behaviour. That is due to the assumption that the doer bears 'psychological disutility' when he misses the prespecified target. This can be interpreted as guilt that one feels when failing to follow well-intended plans, arising perhaps from loss aversion or a desire to act in an internally consistent way. Such self-regulation mechanisms have been studied extensively by psychologists (Heath, Larrick, and Wu, 1999; Locke and Latham, 1990, 2002).

Specifically, denote the savings target chosen by the planner for period t by $\hat{a}_t$. In a model with one asset, a savings target is equivalent to a consumption target $\hat{c}_t \equiv y_t - \hat{a}_t$. Supposing that this sensation of psychological disutility is increasing and convex in the amount by which the target was missed, but produces no utility from exceeding the target, the instantaneous utility function takes the following general form:

$$U(c_t, \hat{c}_t) = u(c_t) - V(c_t, \hat{c}_t)$$

with:

$$V(c_t, \hat{c}_t) = \begin{cases} \psi \, v\big(u(c_t) - u(\hat{c}_t)\big) & \text{for } c_t \geq \hat{c}_t \\ 0 & \text{for } c_t < \hat{c}_t \end{cases}$$

where $u(\cdot)$ is strictly increasing and strictly concave, $v(\cdot)$ is strictly increasing and strictly convex with $v(0) = 0$ and $v'(0) = 0$, and $\psi > 0$. While $u(c_t)$ is the standard utility from consumption, $V(c_t, \hat{c}_t)$ captures the psychological disutility associated with over-consuming relative to the target. In line with the literature, the deviation from the target is expressed in terms of utils.[3] For a given deviation from the target, the disutility is also scaled by a 'self-control' parameter ψ.[4]

[3] In the appendix, I analyse an alternative formulation where the psychological disutility is a function of the monetary amount by which the target is missed.

[4] The above functional form is reminiscent of the regret theory of Loomes and Sugden (1982), although it captures guilt from not following a prespecified plan, rather than regret of choosing an action that turned out to be suboptimal ex post. Relative to a piecewise-linear gain-loss formulation (Koch and Nafziger, 2011, 2016), the above assumes a convex penalty for missing the target, but no benefit from exceeding the target. This approach is analytically convenient, because the disutility from missing the target can enter the doer's problem in a case-independent form, resulting in a smooth objective. Moreover, the above formulation allows to nest the behaviour of a 'classical' agent who follows the devised plan exactly but does not derive any additional utility from exceeding the target.

In exerting influence over the doers' behaviour, the objective of the planner is to maximise the lifetime utility. Naturally, the planner's decision should depend on her beliefs about the doer's strategy, denoted $c_t(\hat{c}_t)$. To focus attention, consider a 'short-run-perfect' Nash equilibrium (SR-perfect NE), in which the doer who is active at time t responds optimally to the target $\hat{c}_t$ chosen by the planner and the planner chooses her targets optimally, correctly anticipating the doers' responses (Fudenberg and Levine, 2006). Thus, in any SR-perfect NE, the doer chooses current consumption c_t to maximise the instantaneous utility $U(c_t, \hat{c}_t)$, subject to the no-borrowing constraint $a_t \geq 0$ (or, equivalently, $c_t \leq y_t$) and taking the target $\hat{c}_t$ as given.

To explicitly account for the difficulty of planning for retirement, I additionally assume that coming up with the consumption-savings targets imposes a cognitive cost on the agent. Similarly to Reis (2006) and Choukhmane (2021), suppose that planning for retirement imposes a fixed cost at the time when the targets are first set. Thus, the cognitive costs of planning arise before the agent starts saving and are independent of the resulting plan. This is consistent with the idea that planning is a prerequisite to financial decision-making and that the mental effort involved is independent of the conclusion ultimately reached. Put plainly, it is just as difficult to figure out that one's desired savings rate should be 5%, 10%, or 15%. Potential determinants of the level of cognitive costs include clarity of the design of the pension system, compulsory financial education, as well as individual numerical ability and stress-related cognitive load. Under an alternative interpretation, the cost of planning may reflect the combination of a monetary cost of seeking financial advice and the cognitive cost of assessing its quality.[5]

Denote the time at which the planner plans for retirement by $\tau \in \{1, 2, ..., T-1, \varnothing\}$, with $\tau = \varnothing$ illustrating non-planning. In periods $t < \tau$, or if $\tau = \varnothing$, the planner does not try to exert influence over the doer, who consumes the entire disposable income y_t. Formally, 'no target' is modelled by setting $\hat{c}_t = y_t$. It is hopefully clear that in this case $U(c_t, y_t)$ is maximised at $c_t = y_t$.

Otherwise, the doer's choice of consumption and savings reflects the trade-off between the utility from immediate consumption and the disutility from missing the target, embedded in maximisation of $U(c_t, \hat{c}_t)$. Taking into account the doers' responses to her targets, the planner chooses the targets optimally, i.e. to smooth consumption across periods $\tau, \tau+1, ..., T$. In a deterministic environment considered in this section, the planner is able to set the optimal targets for the remainder of the life cycle, so that the cost of planning is borne only once. In the final period, the retired agent optimally consumes the entire pension pot of size A_T.

Thus, for each $\tau \neq \varnothing$, the lifetime utility that can be attained by the planner in a SR-perfect NE is given by:

[5] As it will become clearer in section 3, I do not assume that it is always less cognitively costly to save nothing as opposed to a positive amount, but rather that it is always less cognitively costly to adhere to a current default consumption path.

$$\Upsilon(\tau) \;\equiv\; \sum_{t=1}^{\tau-1} \delta^{t-1} u(y_t) + \max_{\hat{c}_\tau, \hat{c}_{\tau+1}, \ldots, \hat{c}_{T-1}} \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1} [u(c_t(\hat{c}_t)) - V(c_t(\hat{c}_t), \hat{c}_t)] + \delta^{T-1} u(A_T) \Big\} - \delta^{\tau-1}\phi$$

where $\delta < 1$ is the (time-consistent) planner's discount factor and $\phi \geq 0$ represents the fixed cognitive cost of planning.

The planner may also forego planning for retirement altogether, in which case write:

$$\Upsilon(\varnothing) \;\equiv\; \sum_{t=1}^{T-1} \delta^{t-1} u(y_t) \;+\; \delta^{T-1} u(y_T)$$

Importantly, within each period preferences of the planner and the doer, represented by $U(c_t, \hat{c}_t)$, coincide. Interpreting $\Upsilon(\tau)$ as the normative measure of the agent's welfare implies that the costs of planning and exercising self-control are treated as welfare-relevant.[6]

2.2 Characterisation of behaviour

The model is solved backwards. That is, I first consider the impact of the targets on the doer's savings behaviour and then I provide a characterisation of the planner's optimal strategy, although chronologically speaking *planning* precedes *saving*.

Doer ('saving'). In every period t, the doer's choice maximises $U(c_t, \hat{c}_t)$. The solution necessarily lies in the interval $c_t(\hat{c}_t) \in [\hat{c}_t, y_t]$. That is because the doer has no incentive to consume less than $\hat{c}_t$ and, at the other extreme, is not able to consume more than y_t due to the no-borrowing constraint. Interior solutions satisfy the following first-order condition:

$$\psi\, v'\big(u(c_t) - u(\hat{c}_t)\big)\, u'(c_t) \;=\; u'(c_t)$$

$$\iff \quad \psi\, v'\big(u(c_t) - u(\hat{c}_t)\big) \;=\; 1 \tag{1}$$

On the other hand, if:

$$\psi\, v'\big(u(y_t) - u(\hat{c}_t)\big) \;<\; 1 \tag{2}$$

then the no-borrowing constraint is binding at the doer's optimum and $c_t(\hat{c}_t) = y_t$. Since $v'(0) = 0$, lax targets close to y_t are ignored by the doer. As the left-hand side of the inequality (2) is decreasing in $\hat{c}_t$, agents with poor self-control (low ψ) need stricter targets (low $\hat{c}_t$) to induce positive savings.

[6] Goldin and Reck (2020) distinguish between different classes of models according to their implications regarding the normativity of the decision-making costs.

Focusing on interior solutions, the first-order condition (1) has intuitive implications. First, from strict convexity of $v(\cdot)$, an agent characterised by a greater self-control parameter ψ is more obedient in the sense that the doer's choice is closer to the planner's target, i.e. $c_t(\hat{c}_t)$ is strictly decreasing in ψ. Moreover, as ψ grows unboundedly, $c_t(\hat{c}_t)$ converges to $\hat{c}_t$, because $v'(\cdot)$ is strictly increasing and $v'(0) = 0$. Second, for a given degree of self-control, a more stringent target makes the doer save more, i.e. $c_t(\hat{c}_t)$ is strictly increasing in $\hat{c}_t$.

In order for the model of costly self-control to have desired properties, it is not sufficient to establish that $c_t(\hat{c}_t)$ is strictly decreasing in ψ. After all, the total cost of self-control given by:

$$\psi\, v\Big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\Big)$$

might still be either increasing or decreasing in ψ. I now derive a condition that ensures that the total cost of self-control associated with a given target is indeed *decreasing* in the self-control parameter ψ. This will allow the model to nest the behaviour of a 'classical' agent, who can implement a desired target without any cost of self-control.

$$\frac{d\,\psi\, v\Big(u(c_t(\hat{c}_t))-u(\hat{c}_t)\Big)}{d\,\psi} \;=\; v\Big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\Big) \;+\; \psi\, v'\Big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\Big)\, u'(c_t(\hat{c}_t))\, \frac{dc_t(\hat{c}_t)}{d\psi} \;<0$$

$$\iff \quad \frac{dc_t(\hat{c}_t)}{d\psi} \;<\; \frac{-v\Big(u(c_t(\hat{c}_t))-u(\hat{c}_t)\Big)}{\psi\, v'\Big(u(c_t(\hat{c}_t))-u(\hat{c}_t)\Big)\, u'(c_t(\hat{c}_t))}$$

That is, for the total cost of self-control to be decreasing in the self-control parameter ψ, the doer's choice $c_t(\hat{c}_t)$ needs to be converging to the target $\hat{c}_t$ 'fast enough' when ψ increases.

To establish under what conditions the above inequality is true, I apply the implicit function theorem (de la Fuente, 2000, Theorem 2.1) to write:

$$\frac{dc_t(\hat{c}_t)}{d\psi} \;=\; \frac{-v'\Big(u(c_t)-u(\hat{c}_t)\Big)}{\psi\, v''\Big(u(c_t)-u(\hat{c}_t)\Big)\, u'(c_t(\hat{c}_t))}$$

Then, the desired inequality is satisfied when:

$$\frac{-v'\Big(u(c_t)-u(\hat{c}_t)\Big)}{\psi\, v''\Big(u(c_t)-u(\hat{c}_t)\Big)\, u'(c_t(\hat{c}_t))} \;<\; \frac{-v\Big(u(c_t(\hat{c}_t))-u(\hat{c}_t)\Big)}{\psi\, v'\Big(u(c_t(\hat{c}_t))-u(\hat{c}_t)\Big)\, u'(c_t(\hat{c}_t))}$$

$$\Longleftarrow \quad \frac{v'(\cdot)}{v''(\cdot)} \;>\; \frac{v(\cdot)}{v'(\cdot)}$$

This leads to the following disciplining assumption.

Assumption 1: *For all non-negative arguments $v(\cdot)$ is convex but not logarithmically convex.*

For example, if one wishes to suppose a power function formulation $v(x) = x^\gamma$, then Assumption 1 is equivalent to $\gamma > 1$, i.e. strict convexity. Such non-linearities are consistent with the model of Fudenberg and Levine (2006), but not with a linear cost of resisting temptation considered by Gul and Pesendorfer (2001).

Planner ('planning') The planner's problem can be represented as a problem of choosing the optimal planning horizon, conditional on the savings targets being chosen optimally for any horizon. Then, the lifetime utility attained by the planner in a SR-perfect NE is:

$$\Upsilon^* \equiv \max_{\tau \in \{1,2,\ldots,T-1,\varnothing\}} \Upsilon(\tau)$$

For any period τ when the planning starts, the problem of choosing the optimal targets is:

$$\max_{\hat{c}_\tau, \hat{c}_{\tau+1}, \ldots, \hat{c}_{T-1}} \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1} [u(c_t(\hat{c}_t)) - \psi\, v\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)] + \delta^{T-1} u(A_T) \Big\}$$

where $A_T = \sum_{t=\tau}^{T-1} R^{T-t}(y_t - c_t(\hat{c}_t)) + y_T$.

Recall that the planner needs to select a strict enough target in order to induce positive savings. From inequality (2), the minimum target required to induce positive savings $\underline{\hat{c}}_t$ satisfies:

$$\psi\, v'\big(u(y_t) - u(\underline{\hat{c}}_t)\big) = 1 \tag{3}$$

This implies that the agents characterised by poor self-control (low ψ) require a stricter minimum target (low $\underline{\hat{c}}_t$) to be able to save any positive amount. Since this minimum target depends on y_t, it is indexed by t.

Notice that it is never optimal for the planner to set an 'ineffective strict target' $\hat{c}_t \in [\underline{\hat{c}}_t, y_t)$. Such targets result in zero savings, but nevertheless impose a strictly negative penalty for missing the target. Furthermore, when the target is chosen from this interval, the doer's response is not captured by the first-order condition (1), but rather by the binding no-borrowing constraint, which is important to keep in mind when solving the planner's problem.

Because an 'optimal lax target' $\hat{c}_t = y_t$, which allows the doer to consume his entire disposable income, may be desirable in certain periods, it will be useful to distinguish explicitly between the periods in which the planner optimally induces positive savings and the ones in which she does not (conditional on planning). Thus write the planner's problem conditional on a particular planning horizon as:

$$\max_{\hat{c}_\tau, \hat{c}_{\tau+1}, \ldots, \hat{c}_{T-1}} \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1} [u(c_t(\hat{c}_t)) - \psi\, v\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)] + \delta^{T-1} u(A_T) \Big\}, \text{ subject to:}$$

1. $A_T = \sum_{t=\tau}^{T-1} R^{T-t}(y_t - c_t(\hat{c}_t)) + y_T$

2. $\hat{c}_t \leq \underline{\hat{c}}_t$

which can be represented using the following Lagrangian:

$$\mathcal{L}_\tau = \{\sum_{t=\tau}^{T-1} \delta^{t-1}[u(c_t(\hat{c}_t)) - \psi\, v\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)] + \delta^{T-1} u(\sum_{t=\tau}^{T-1} R^{T-t}(y_t - c_t(\hat{c}_t)) + y_T)\} + \sum_{t=\tau}^{T-1} \lambda_\tau[\underline{\hat{c}}_t - \hat{c}_t]$$

For any period $t = \tau, \tau+1, ..., T-1$, the first-order condition associated with $\hat{c}_t$ is:

$$\delta^{t-1}[u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - \psi\, v'\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)\,(u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - u'(\hat{c}_t)] + \delta^{T-1}u'(A_T)R^{T-t}(-\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}) - \lambda_t \;=\; 0$$

$$\iff \quad u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - \psi\, v'\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)\,(u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - u'(\hat{c}_t)) \;=\; (\delta R)^{T-t}u'(A_T)\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} + \tfrac{\lambda_t}{\delta^{t-1}}$$

Taking into account the doer's optimal response to the target, captured by (1), the above simplifies to:

$$u'(\hat{c}_t) \;=\; (\delta R)^{T-t}u'(A_T)\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} + \tfrac{\lambda_t}{\delta^{t-1}}$$

The first-order condition associated with λ_t is simply:

$$\underline{\hat{c}}_t - \hat{c}_t \;=\; 0 \iff \hat{c}_t = \underline{\hat{c}}_t$$

By the Kuhn-Tucker Theorem (de la Fuente, 2000, Theorem 1.18), if the solution has $\lambda_t > 0$, then the constraint is indeed binding at the optimum. In this case, the planner optimally sets the optimal lax target $\hat{c}_t = y_t$. In other words, if the target $\underline{\hat{c}}_t$ is optimal in the range $\hat{c}_t \leq \underline{\hat{c}}_t$, then $\hat{c}_t = y_t$ must be maximising the planner's objective.[7]

It follows that in periods when:

$$u'(\underline{\hat{c}}_t) \;>\; (\delta R)^{T-t}u'(A_T)\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}|_{\hat{c}_t=\underline{\hat{c}}_t}$$

it is optimal for the planner to set $\hat{c}_t = y_t$ and not induce any positive savings. We can further simplify this expression by again applying the implicit function theorem:

[7]When the planner's objective function is pseudo-concave and the constraints are quasi-concave, the first-order conditions outlined above are both necessary and sufficient for an optimum. If, in addition, the objective function is strictly quasi-concave, then this optimum is unique (de la Fuente (2000), Theorem 1.19). In the planner's problem, the constraints are linear in $\hat{c}_t$ and therefore quasi-concave. Given the doer's optimal response, the psychological disutility is fixed and thus the planner's objective function is just a discounted sum of strictly concave instantaneous utility functions, which is itself strictly concave and therefore strictly quasi-concave. Thus, the above first-order conditions identify a unique optimum.

$$\frac{dc_t(\hat{c}_t)}{d\hat{c}_t} \;=\; \frac{-\psi\, v''\Big(u(c_t)-u(\hat{c}_t)\Big)\,(-u'(\hat{c}_t))}{\psi\, v''\Big(u(c_t)-u(\hat{c}_t)\Big)\,u'(c_t(\hat{c}_t))} \;=\; \frac{u'(\hat{c}_t)}{u'(c_t(\hat{c}_t))} \;\geq\; 1$$

which is decreasing in ψ, because $c_t(\hat{c}_t)$ is decreasing in ψ and converges to $\hat{c}_t$. Then, keeping in mind that $c_t(\underline{\hat{c}}_t) = y_t$, we can see that the planner sets the optimal lax targets in periods when:

$$u'(c_t(\underline{\hat{c}}_t)) = u'(y_t) \;>\; (\delta R)^{T-t} u'(A_T) \tag{4}$$

just as a standard life-cycle savings model with borrowing constraints would prescribe.

On the other hand, if the solution returns $\lambda_t \leq 0$, then the constraint $\hat{c}_t \leq \underline{\hat{c}}_t$ is slack at the optimum found in the range $\hat{c}_t \leq \underline{\hat{c}}_t$. In this case, set $\lambda_t = 0$ in the Lagrangian to not impose the constraint to be binding and find the optimal interior target. However, because the doer's policy function is not differentiable at $\hat{c}_t = \underline{\hat{c}}_t$, it cannot be ruled out that $\hat{c}_t = y_t$ maximises the planner's objective, even though the Lagrangian has an interior solution. To rule out this possibility, one should compare the lifetime utilities resulting from choosing a 'seemingly optimal' interior target and the optimal lax target. Intuitively, the interior target would be globally optimal if the utility gain from consumption smoothing exceeds the cost of exercising self-control in period t.

In periods with $\lambda_t = 0$, the optimal interior target satisfies the standard Euler equation:

$$u'(c_t(\hat{c}_t)) \;=\; (\delta R)^{T-t} u'(A_T) \tag{5}$$

That is, given the doer's response, the optimal target equalises the marginal utility from consumption in period t with the marginal utility from retirement savings and the usual considerations that have implications for the optimal shape of the saving schedule, such as wage growth over time and the trade-off between returns and impatience, apply.

The conditions (4) and (5) suggest that conditional on the planning horizon and the size of the pension pot A_T, the agent's self-control does not affect her optimal savings in period t. Why is that? Notice that the doer's optimal response, captured by (1), implies that the cost of exercising self-control is identical across all effective targets. Then, the planner selects the target that is optimal from the perspective of pure consumption smoothing.

So what is the impact of self-control on the optimal targets and pension wealth accumulation? First, notice that inducing positive savings is more costly for agents characterised by poor self-control. That is because they need to set a stricter target to induce a given level of savings ($c_t(\hat{c}_t)$ is decreasing in ψ) and, under Assumption 1, they face a greater cost of self-control associated with a particular target ($\psi\, v\big(u(c_t(\hat{c}_t)) - u(\hat{c}_t)\big)$ is also decreasing in ψ). Conversely, agents with better self-control bear a lower cost of inducing positive savings, which makes them less likely to set the lax targets. All else held constant, if an agent with self-control parameter $\tilde{\psi}$ saves a positive amount in period t, so do all agents with better self-control $\psi > \tilde{\psi}$. As a result, agents with better self-control spread their retirement savings over a greater number of periods.

Second, in periods with positive savings, the Euler equation (5) applies independently of the degree of self-control. Since agents with poor self-control are saving over fewer periods, the Euler equation implies that they actually save more in those periods when they set themselves effective targets, because of the greater marginal utility from saving. However, since in a standard consumption-savings problem with borrowing constraints increasing the number of periods in which saving can take place can only lead to higher wealth accumulation, we have the following result:

Lemma 2.1 (concentrated under-saving): *For any exogenous planning horizon $\tau \neq \varnothing$, the number of periods in which an agent sets effective savings targets $\hat{c}_t < \underline{\hat{c}}_t$ as well as the resulting retirement wealth A_T are (weakly) increasing in the self-control parameter ψ. In comparison, savings of an agent with lower ψ are concentrated in fewer periods with positive savings, during which she saves more than her counterparts with better self-control.*

The proofs of all results are provided in the appendix.

Next, how does the planner decide *when* to implement a savings plan, if at all? In a deterministic decision environment, the planner is able to calculate the optimal sequence of targets, and the resulting lifetime utility, for any planning horizon already in period 1. Then, the choice of the optimal horizon boils down to a simple cost-benefit analysis. While implementing a savings plan earlier on allows the agent to achieve a higher degree of consumption smoothing, it also imposes a cognitive cost of a greater present value. To break possible ties in the definition of Υ^*, suppose that when indifferent, the planner chooses an option with a lower present value of the cognitive cost.

Such cost-benefit approach to choosing the planning horizon seems natural at first, but it poses a serious conceptual challenge - the infinite regress problem (Conlisk, 1996; Lipman, 1991). To illustrate, suppose that solving a complicated optimisation problem is cognitively costly, but yields some benefit. The issue is that the decision about whether or not to bear the cognitive cost in the first place only adds a further layer of complexity to the original optimisation problem. Analogously, the decision about the decision about bearing the cognitive cost is even more complex, and so on. In a recent paper, however, Alaoui and Penta (2021) derive conditions under which the decision to reason can be represented by an 'as-if' cost-benefit analysis. Intuitively, they find that this approach is justified in a class of problems where reasoning is instrumental to the agent's ultimate choice, her reluctance to reason is identical across problems that differ by a constant payoff (whenever reasoning is not instrumental), and she does not have a strict preference for committing to ignore what she learns. Since these conditions appear quite weak, Alaoui and Penta (2021) argue that the cost-benefit approach is justifiable in many contexts.[8]

[8] At the same time, the conditions derived by Alaoui and Penta (2021) are inconsistent with certain psychological effects, such as thinking aversion (when thinking may be detrimental), rumination, and choking under

Proceeding with the analysis, from the definition of $\Upsilon(\tau)$ it should be clear that if the cognitive cost ϕ is zero, then the planner sets up her savings targets in the very beginning of the life cycle, as this puts the fewest constraints on the planner's problem. When $\phi > 0$, we can compare two adjacent periods in order to illustrate when the planner postpones saving for retirement:

$$\Upsilon(\tau+1) \;\geq\; \Upsilon(\tau) \;\iff\; (1-\delta)\phi \;\geq\; \Big\{\max_{\hat{c}_\tau,\hat{c}_{\tau+1},\ldots,\hat{c}_{T-1}}\big\{\sum_{t=\tau}^{T-1}\delta^{t-\tau}[\max_{c_t\leq y_t}u(c_t)-V(c_t,\hat{c}_t)]+$$

$$\delta^{T-\tau}u(A_T)\big\}\Big\} - \Big\{u(y_\tau)-\max_{\hat{c}_{\tau+1},\ldots,\hat{c}_{T-1}}\big\{\sum_{t=\tau+1}^{T-1}\delta^{t-\tau}[\max_{c_t\leq y_t}u(c_t)-V(c_t,\hat{c}_t)]+\delta^{T-\tau}u(A_T)\big\}\Big\}$$

That is, the postponement of retirement planning is desirable if bearing the discounted cognitive cost of decision-making outweighs the utility loss from smoothing consumption over a shorter horizon, represented by the right-hand side of the inequality.

Under a regularity assumption spelled out below, the utility loss from smoothing consumption over an arbitrarily shorter horizon is monotonically increasing in self-control and, in particular, is maximised for the agents with perfect self-control ($\psi \to \infty$), but always equal to zero for agents with extremely poor self-control ($\psi \to 0$). A natural result follows that, other things being equal, poor self-control leads to delays in planning and saving for retirement. Such delays reflect the interaction between the two features of the model. With zero cognitive cost, the degree of self-control would have no impact on timing of retirement planning.

Before stating the result, we need to introduce the following:

Assumption 2: *Disposable income is non-decreasing during the working life, i.e.* $y_t \leq y_{t+1}$ *for all* $t = 1, 2, T-2$*, and* $(\delta R) \leq 1$.

When Assumption 2 is satisfied, the rationale for saving for retirement is the strongest in the last period of the working life, and weakens monotonically in the distance from retirement. That is because according to the Euler equation (5), the marginal improvement to consumption utility, which reflects the difference between marginal utility from consuming more at T and marginal disutility from consuming less at t, is the greater, the nearer the agent is to retirement. This assumption results in incentives to save which are consistent with a typical life-cycle savings path (see, e.g., Browning and Lusardi, 1996; Scholz, Seshadri, and Khitatrakun, 2006).

The role of Assumption 2 in the proof of Lemma 2.2 is to rule out parametrisations for which the marginal improvement to consumption utility from extending the planning horizon is sometimes greater for agents with poorer self-control. This could happen if, for example, one of the periods in mid-life $t < T-1$ provided the agent with an exceptionally high disposable income. Then, extending the planning horizon to include t might improve the utility attained

pressure. Arguably, these are not of first-order importance when exerting mental effort to formulate a financial plan.

by an agent with poor self-control by a greater margin than the utility attained by an otherwise identical agent with better self-control, if t is the only period in which the agent with poor self-control saves a positive amount, but the positive savings of the agent with better self-control are spread over multiple later periods.

When such cases are ruled out, we have the following result, stating that the agents with better self-control and lower cognitive costs start planning for retirement earlier on during their working lives. Note that this is the only result that relies on Assumption 2 being satisfied.

Lemma 2.2 (planning horizon): *Suppose that Assumption 2 holds. Let $\tau^* \neq \varnothing$ correspond to the planning horizon that maximises the planner's utility. Then, τ^* is (weakly) decreasing, i.e. the optimal planning horizon is longer, in self-control ψ and (weakly) increasing, i.e. the optimal planning horizon is shorter, in cognitive cost ϕ.*

Next, when does the planner forego retirement saving altogether, i.e. displays complete inertia in planning for retirement? It seems intuitive that there should exist a threshold value for the cognitive costs, above which the cost outweighs the utility benefits from retirement saving and the planner optimally sets no targets for the doers. To make this point more precise, note that when evaluated at $\phi = 0$, the lifetime utilities associated with each planning horizon can be clearly ordered:

$$\Upsilon(\varnothing) \ \leq \Upsilon(T-1) \ \leq \ ... \ \leq \Upsilon(1) \quad \text{for } \phi = 0$$

Furthermore, the impact of an increase in cognitive costs varies with the planning horizon:

$$\frac{d\,\Upsilon(\varnothing)}{d\,\phi} = 0 \ > \ \frac{d\,\Upsilon(T-1)}{d\,\phi} = -\delta^{T-2} \ > \ ... > \ \frac{\Upsilon(1)}{d\,\phi} = -1$$

Together with the observation that each $\Upsilon(\tau)$ is increasing in the self-control parameter ψ, these regularities imply:

Proposition 2.1 (non-saving): *There exists a threshold $\overline{\phi}(\psi) > 0$, such that for $\phi \geq \overline{\phi}(\psi)$ no planning for retirement takes place, i.e. $\Upsilon^* = \Upsilon(\varnothing)$. Moreover, $\overline{\phi}(\psi)$ is increasing in ψ.*

The above result formalises the intuition that the planner does not engage in retirement preparation, leading to a corner solution in optimal savings, when the cognitive costs of planning outweigh the potential utility gains from consumption smoothing. Since those gains are lower for agents with poor self-control, the doer's propensity to follow a devised savings target informs the planner's decision about whether to set the target in the first place. When the two channels that might lead to undersaving are separated, their impacts appear to be substitutable. Agents characterised by outstanding self-control would need to find planning extremely burdensome to be prevented from saving. On the other hand, agents with poor self-control would not plan for retirement even if the associated cognitive costs were small.

Combining the previous three results we can conclude that poor self-control and high cognitive costs of planning unambiguously lead to lower levels of consumption in retirement. That reflects the joint impact of under-saving over a fixed horizon (Lemma 2.1), a tendency to plan over shorter horizons (Lemma 2.2), and a propensity to not undertake any planning at all (Proposition 2.1).

Corollary 2.1 (under-consumption in retirement): *Retirement consumption A_T^* supported by the optimal saving schedule implemented by the planner is (weakly) increasing in self-control ψ and (weakly) decreasing in cognitive costs of planning ϕ.*

In addition to the agent's self-control and cognitive costs, her level of income is also an important input into the planner's decision. Consider the following comparative statics exercise, whereby the vector representing the agent's disposable income profile $(y_1, ..., y_T)$ is multiplied by a constant factor $\mu > 1$. In a standard life-cycle model with an isoelastic (CRRA) utility function, such scaling up has no impact on the optimal savings rate in any period, as it leaves all the optimality conditions unaffected. In the model with cognitive costs of planning, in contrast, it can affect the planner's decision about the optimal planning horizon, because the differences in absolute utility levels that can be attained over various planning horizons are inflated. In particular, following the logic underlying Lemma 2.2, scaling up the income profile makes planning over shorter horizons, as well as not planning at all, less desirable for a given cost of planning ϕ. Furthermore, these adjustments to the optimal planning horizon can only increase wealth accumulation of an agent. Thus, when the planning horizon is extended, the wealth-to-income ratio (defined as a ratio of optimal pension wealth to average, or aggregate, working life disposable income) can increase in the scaling parameter μ.

Corollary 2.2 (disposable income and wealth-to-income ratio): *Suppose that the utility function $u(\cdot)$ takes the CRRA form. For an agent with either positive costs of planning ($\phi > 0$) or imperfect self-control ($\psi < +\infty$), the wealth-to-income ratio attained at the planner's optimum is (weakly) increasing in the scale of her income profile μ.*

Consistent with the empirical evidence indicating that low-income households save disproportionately little (e.g. Dynan et al., 2004), the model generates a link between the agent's level of income and the propensity to accumulate wealth, which is missing from the classical consumption-savings theory. Assuming that self-control and cognitive costs are uncorrelated with income, agents with higher incomes are more likely to plan over long horizons and have higher wealth-to-income ratios as a result.[9]

[9] Supposing instead that income and self-control are positively correlated (e.g. due to stress-related cognitive load) and that income and cognitive costs of planning are negatively correlated (e.g. due to the mediating role of education) would only strengthen the relationship between the level of income and financial preparedness for retirement.

Taken together, Proposition 2.1 and Corollary 2.2 indicate that 'non-savers', i.e. individuals who fail to save privately for retirement, are characterised by poor financial self-control, high cognitive costs, and/or low incomes. This appears to be a realistic prediction. As an example, consider the 'Attitudes to Pensions' survey conducted by the UK government in 2012 to inform extensive reforms of the pension system. The survey classified 19% of respondents as non-savers. These individuals had no sources of private pension wealth, including workplace pensions, saving accounts, and property that a respondent would consider selling or downsizing to finance their retirement. Compared to the rest of the sample, those individuals were more likely to come from a low-income household and to exhibit symptoms of poor financial self-control and high cognitive costs of financial planning. Importantly, the respondents appear (at least somewhat) aware of their self-control and decision-making skills, which is consistent with the notion of *rational inaction* outlined in Proposition 2.1.[10]

Furthermore, the above results identify two novel inputs into the endogenous decision to formulate a savings plan and when to implement it, if at all. These are the agent's propensity to follow through with a plan (i.e. self-control) and her level of income. By identifying potential drivers of non-planning, the model contributes to the empirical literature on planning, financial knowledge, and wealth accumulation (see, e.g., Ameriks et al., 2003; Hurd and Rohwedder, 2013; Lusardi and Mitchell, 2007, 2011; Van Rooij et al., 2012). This literature has documented that non-planning is widespread and associated with substantially lower wealth accumulation, but remains largely silent on the determinants of non-planning.[11] The theoretical model proposed above constitutes a step towards filling this gap. In addition, the model accounts for the dynamics of planning and wealth accumulation (see Lemma 2.1 and Lemma 2.2), which provides some key intuitions underlying the relationship between current wealth and past planning decisions observed in the cross-sectional data.

[10]Specifically, while 37% of households from the lowest income category (annual household income of £12,000 or less) had no private pension wealth, this proportion was equal to 3% in the highest income category (annual household income of £44,000 or more). A similarly strong link has been observed across education categories. In relation to financial self-control, 29% of non-savers (compared to 59% in the subsample of savers) said that they were keeping up with their bills and credit commitments without any difficulties, 35% (58%) reported that they were putting some money aside for emergency situations. Regarding the mental effort associated with financial decisions, 41% of non-savers (12% of savers) said that 'they would have no idea about what they needed to do when making important financial decisions, such as taking out a mortgage, loan, or pension', 37% (16%) reported their financial knowledge to be poor, and 84% (51%) stated that they had no idea what their retirement income would be (MacLeod, Fitzpatrick, Hamlyn, Jones, Kinver, and Page, 2012). Of course, such bivariate relationships ignore common determinants (e.g. age), but unfortunately the descriptive DWP report does not contain multivariate regression results.

[11]For example, in the sample of Lusardi and Mitchell (2011), only 43% of American adults say that they have ever tried to figure out what they need to finance their retirement, including only 57% of those aged 50-65. Similar results are found in other developed countries (Keane and Thorp, 2016).

To conclude the analysis in this section, observe that the presence of costly planning and self-control creates a particular form of non-fungibility (imperfect substitutability) between public pensions and privately accumulated pension wealth. Specifically, consider a small exogenous increase in y_T. For a classical agent, such a change results in a proportional decrease in retirement savings in all periods with positive savings. For an agent with either positive costs of planning or exercising self-control, in contrast, the decrease in private savings can be 'more than proportional' if a reduced incentive to save results in a shorter planning horizon or, at the extreme, in complete non-planning. Furthermore, the agent is the more likely to shorten her planning horizon, the poorer her self-control, which again points to non-trivial interactions between self-control and costly planning.

Corollary 2.3 (non-fungibility of public and private pension wealth): *An exogenous increase in y_T results in a disproportionate reduction in pension wealth A_T^* whenever it shortens the optimal planning horizon of an agent with either positive costs of planning ($\phi > 0$) or imperfect self-control ($\psi < +\infty$). Additionally, when Assumption 2 holds, for any planning horizon it is the agents with poor self-control who are most likely to shorten their optimal planning horizon following an increase in y_T.*

An interesting implication of Corollary 2.3 is the possibility that retirement *consumption* A_T^* implemented by the planner might decrease with y_T. That is the case whenever a marginal increase in public pension benefit y_T makes an agent shorten her planning horizon and thus accumulate disproportionately lower private pension wealth. Because this effect requires the cost of either planning or self-control to be non-trivial, it is absent from the standard theory. In other words, a classical agent would never reduce her retirement consumption when the pension benefit is raised.

2.3 Comparison with alternative theories

Having derived the main predictions of the model, I compare it with the most prominent alternatives - the classical consumption-savings model, the two-system model without cognitive costs of planning, and the model of procrastination due to naiveté about self-control by O'Donoghue and Rabin (1999a, 2001). I argue that explicitly accounting for, and separating, costly planning and imperfect self-control results in a framework that captures a range of empirically-relevant behaviours and is thus well-suited to analysing the impact of specific policies.

Classical theories. As discussed above, a classical agent who has no self-control problems and can plan for retirement effortlessly is nested within my framework. The introduction of behavioural features into the model results in two additional degrees of freedom, thus calling for justification. In reality, undersaving for retirement is seen as an important policy issue, reflecting the fact that many workers seem to accumulate insufficient private pension wealth (OECD, 2019). The classical theory is in principle able to predict that a mass of agents will have exactly zero savings, but this would require some unrealistic assumptions. A classical agent would not save if either she has a preference for borrowing against future retirement income, or saving imposes high transaction costs. Given the shape of a typical life-cycle income profile, the former explanation would require an extremely low discount factor (Browning and Lusardi, 1996; Scholz et al., 2006). The latter, in turn, would require unreasonably high transaction costs (O'Donoghue and Rabin, 1999b). Thus the two mechanisms have little appeal in explaining a seeming lack of preparedness for retirement. Without imposing prohibitively high transaction costs, the classical model is also unable to replicate the stylized fact that low-income households save disproportionately little (Dynan et al., 2004).

Furthermore, the classical model cannot account for undersaving understood as the discrepancy between one's long-term goals and actual behaviour. However, the prevalence of undersaving in this sense is evident in the observed demand for commitment (Ashraf, Karlan, and Yin, 2006; John, 2020) and survey responses according to which over two thirds of individuals find their own savings rates 'too low' (Choi, Laibson, Madrian, and Metrick, 2002). Of course, the classical theory rules out usefulness of any paternalistic policies or nudging, even among individuals with very little savings.

Two-system model. I modify the standard two-system framework by introducing the cognitive cost of planning. For two reasons, this is potentially useful. First, rich empirical evidence reviewed above documents the role of planning in wealth accumulation, raising a natural question about the determinants of non-planning, and optimal planning horizon more broadly. Accounting for the cognitive cost of planning thus allows to reconcile the theory with the empirical observations, as well as formally analyse the interplay between self-control and planning in generating specific behavioural patterns. Second, the current model emphasizes the importance of corner solutions in retirement savings, which arise due to rational inaction. As it will be demonstrated in section 3, this is crucial for replicating the observed responses to changing the default option via automatic enrolment into private pensions. Thus the model enriched with the notion of costly planning not only allows to tackle novel theoretical questions, but can also improve our understanding of the impact of policies that are increasingly widespread in the domain of retirement savings.

Naïve procrastination. In the model of O'Donoghue and Rabin (1999a, 2001), procrastination arises due to naiveté about one's self-control problem. At any point in time, a naïve agent postpones an unpleasant activity that involves immediate cost and delayed benefits, because he overestimates his willingness to carry out the task in the following period. In contrast, an agent who is sophisticated about his self-control problem might delay the completion of the task for several periods, but eventually completes the task.[12] This mechanism is fundamentally different from a model of rational inaction proposed here, which results in some divergent predictions. To make this point more precise, I propose two ways of drawing a link between the two theories.

First, one might interpret the immediate cost in the naïve procrastination framework as corresponding to the cognitive cost of planning. Then, the model implies that inertia in pension choices relies on naiveté and, in particular, learning one's self-control would eliminate or curb it. In contrast, the model of rational inaction implies that inertia arises due to sophistication about one's (lack of) self-control and, in particular, learning one's self-control would make non-planning more widespread. Lending support to the notion of rational inaction, recall that non-savers from the 'Attitudes to Pensions' survey self-reported their difficulties with keeping up with financial commitments and inability to save for emergencies, thus exhibiting at least partial sophistication. A similar case can be made based on experimental studies that find evidence for demand for costly commitment in the domain of savings (see John, 2020). Importantly, although the two models are in principle capable of generating similar predictions, taking a stand on the underlying model of decision-making is key for comparative statics and determining the desirability of certain policies, e.g. imposing hard deadlines in pension choices or choosing the particulars of a default pension scheme.[13]

Second, under an alternative interpretation, the naïve procrastination framework applies to the second stage of the model analysed above, where having exerted the effort to plan for retirement, the agent is supposed to carry out the plan, e.g. join the workplace pension scheme by filling out a relevant form. In this case, the agent's tendency to procrastinate is treated as one of the inputs into the first-stage decision about planning. Explicitly separating those two steps of financial decisions seems consistent with the evidence from neuroscience indicating that separate parts of the brain are responsible for impulse inhibition (orbitofrontal cortex) and planning (dorsolateral prefrontal cortex) (Camerer, 2013; Duckworth, Milkman, and Laibson, 2018), thus justifying the introduction of an additional stage into the model of decision-making.

[12]Although a partially naïve agent does not always procrastinate, O'Donoghue and Rabin (2001) show that for any degree of naiveté there exists an environment in which he does.

[13]Forbidding delays is unambiguously welfare-improving under procrastination due to naiveté when the current costs of an action are either deterministic or negligibly small. O'Donoghue and Rabin (2001) introduce the model in which the agent decides not only when to complete a task, but also which task to complete. They demonstrate that the agent might procrastinate more on more important tasks, which may seem at odds with the cost-benefit reasoning underlying rational inaction. Note, however, that O'Donoghue and Rabin (2001) define a task to be more important when it is characterised by greater future benefits as well as immediate costs. Such an exercise appears better suited to intra-personal comparisons, e.g. an individual is more likely to procrastinate on choosing a healthcare plan than mowing a lawn. In contrast, result presented in Proposition 2.1 focuses on inter-personal comparisons, i.e. holding the cognitive cost of planning constant, an individual with better self-control is more likely to undertake retirement planning.

2.4 Introducing uncertainty and re-planning

It is interesting to ask how the planner's strategy might change in face of uncertainty, especially in the context of retirement savings. Shocks to investment returns, income, employment, health, and survival, and the expectations thereof, play an important role in shaping the optimal savings plan. In this section, I briefly discuss an extension of the model that allows to study uncertainty and re-planning.

In period t, the state is summarised by a vector containing all realisations of the stochastic variables up to time t (e.g. disposable income or returns) Ω^t, the current size of the agent's pension pot A_t, and the current plan $\hat{C}^t \equiv (\hat{c}_t^t, ..., \hat{c}_{T-1}^t)$. Note that a special case of the current plan is no plan at all, where $\hat{c}_i^t = y_i$ for all $i \in \{t, ..., T-1\}$.

The planner's decision is to whether or not update the current plan and, if yes, how. An advantage of the two-system framework is that the planner has time-consistent preferences and thus her dynamic problem can be formulated recursively. For any period $t < T$, define the expected continuation payoff associated with not updating the plan as:

$$\mathcal{V}_t^0(A_t, \hat{C}^t, \Omega^t) \equiv$$
$$\mathbb{E}_{\Omega^t}\left[(u(c_t(\hat{c}_t^t)) - V(c_t(\hat{c}_t^t), \hat{c}_t^t)) + \delta \max\{\mathcal{V}_{t+1}^0(A_{t+1}, \hat{C}^{t+1}, \Omega^{t+1}); \mathcal{V}_{t+1}^1(A_{t+1}, \hat{C}^{t+1}, \Omega^{t+1})\}\right]$$

while updating the plan yields:

$$\mathcal{V}_t^1(A_t, \hat{C}^t, \Omega^t) \equiv$$
$$\max_{\hat{c}_t^t, \hat{C}^{t+1}} \mathbb{E}_{\Omega^t}\left[(u(c_t(\hat{c}_t^t)) - V(c_t(\hat{c}_t^t), \hat{c}_t^t)) +\right.$$
$$\left.\delta \max\{\mathcal{V}_{t+1}^0(A_{t+1}, \hat{C}^{t+1}, \Omega^{t+1}); \mathcal{V}_{t+1}^1(A_{t+1}, \hat{C}^{t+1}, \Omega^{t+1})\}\right] - \phi$$

where $A_{t+1} = R_{t+1}(A_t + y_t - c_t(\hat{c}_t^t))$ and $\hat{C}^t = (\hat{c}_t^t, \hat{C}^{t+1})$. In particular, this formulation assumes that each instance of active planning imposes a fixed cognitive cost of $\phi > 0$. It also incorporates the planner's (rational) expectation of future re-planning into her current problem.

To complete the definition of the planner's continuation payoff, note that in period T when the agent retires, her payoff is simply:

$$\mathcal{V}_T(A_T, \hat{C}^T, \Omega^T) \equiv u(A_T)$$

Although a full analysis of planning under uncertainty would require further parametric assumptions, in particular regarding the underlying stochastic processes, it remains a reasonable conjecture that agents characterised by high cognitive costs of planning and/or poor self-control should be less likely to set a new plan for any given state. Following several periods of non-planning, this might lead to coming up with some plan eventually and a version of concentrated under-saving, as in Lemma 2.1.

3 Automatic enrolment into private pensions

Automatic enrolment into private pensions, typically offered at a workplace, is an increasingly popular policy tool designed to counteract undersaving (DWP, 2020; OECD, 2021). Under automatic enrolment, an individual has to take a deliberate action to opt out of the scheme, but participates otherwise. When the associated financial incentives are not affected, automatic enrolment constitutes merely a 'nudge', because active decision-makers can easily avoid participation (Sunstein and Thaler, 2008).

Most empirical evidence on the impact of automatic enrolment on participation, contribution rates, and asset allocation comes from firm-level implementations in the US, see Madrian and Shea (2001), Thaler and Benartzi (2004), and Choi et al. (2004) for seminal contributions. Across studies, several regularities are robustly observed. First, automatic enrolment has a substantial positive impact on participation rates. Second, many of those automatically enrolled tend to passively accept the default contribution rate and asset allocation. As a result, the impact of automatic enrolment on total savings remains ambiguous. In particular, automatic enrolment paired with a low default contribution rate might have a negative impact on wealth accumulation by crowding out discretionary savings. Third, because individuals tend to refrain from making active choices, automatic enrolment reduces variation in wealth accumulation across workers. Circumventing the problems with selection into treatment, consistent findings are also reported by Cribb and Emmerson (2020) who analyse the effects of automatic enrolment in the UK, where the policy has been implemented nationwide.

Without a formal model, however, it is difficult to interpret and evaluate the available empirical evidence, a concern voiced in recent reviews of the literature by Beshears et al. (2018a) and Bernheim and Taubinsky (2018). Importantly, what determines whether the agent's total savings increase or decrease as a result of automatic enrolment into a workplace pension? Do higher savings necessarily imply a welfare improvement? In this section, I employ the model with cognitive costs of planning and imperfect self-control to study the effects of automatic enrolment on total savings and welfare. Rational inaction predicted by the model can rationalise the empirical regularities outlined above and provides a lens through which this evidence may be interpreted, thus highlighting the usefulness of explicitly allowing for corner solutions in retirement-related decisions.

Consider a situation in which a policymaker may introduce an automatic enrolment policy at $t = 0$, before an agent can start saving for retirement privately. Let the default pension plan be defined by $(\alpha_1, ..., \alpha_{T-1}, -\beta)$ where $\alpha_t \geq 0$ denotes the contribution collected in period t and $\beta \geq 0$ is the additional pension benefit financed with the contributions. Subsequently, at $t = 1$ the planner decides whether to opt out of the scheme or not. If she does, her problem is identical as in sections 2.1-2.2. If she remains enrolled, she faces a problem of smoothing consumption over the life cycle when the disposable income is given by $\tilde{y}_t = y_t - \alpha_t$ for $t \leq T-1$ and $\tilde{y}_T = y_T + \beta$. Thus, although opting out of the default pension plan is costless, any subsequent active decisions again impose a fixed cognitive cost of planning $\phi > 0$.[14]

In defining the default pension scheme, it will prove useful to relate its parameters to a saving schedule of a classical agent facing a disposable income profile $(y_1, ..., y_{T-1}, y_T)$. Let a_t^C denote the optimal savings of such classical agent in period t, and

$$b^C \equiv \sum_{t=1}^{T-1} R^{T-t} a_t^C$$

be her additional consumption in retirement financed with the working-life savings. Of course, no agent would have any incentive to opt out of a scheme that collects contributions $\alpha_t = a_t^C$ for all $t \leq T-1$ and provides a benefit $\beta = b^C$, irrespective of her self-control and cognitive costs. That is because such a scheme maximises the agent's lifetime utility by allowing her to smooth her consumption perfectly without bearing any costs of self-control or planning.

A more interesting question to ask is: When does the agent accept enrolment into a scheme that deviates from this ideal benchmark? In answering, it is convenient to separate the welfare implications of participation from subsequent decisions about residual savings. Thus, define an 'imperfect pension scheme' as one that leaves no room for subsequent residual savings decisions but does not necessarily maximise the agent's lifetime utility. This is achieved by specifying a default scheme that generates a disposable lifetime income profile which mimics the shape of the classical agent's consumption profile, but provides proportionally lower levels of consumption in every period. Then, an agent who participates in the scheme has no incentive to save privately further, because her default consumption path already satisfies the classical agent's Euler equations.

Definition 1: *An* imperfect pension scheme *collects contributions of*

$$\alpha_t = \gamma a_t^C + (1-\gamma) y_t$$

and provides a pension benefit of

$$\beta = \gamma \Big(\sum_{t=1}^{T-1} R^{T-t} a_t^C \Big) - (1-\gamma) y_T$$

for some $\gamma \in (0, 1]$, effectively allowing the agent to follow the first-best consumption path scaled by γ without bearing any costs of self-control or planning.

[14] Why couldn't the doer opt out instead? I interpret the agent's self-control problem as affecting his propensity to spend out of current disposable income due to short-lived temptations or inattention. The doer, however, takes the disposable income as given and therefore does not undo the commitment made by the planner.

It is straightforward to verify that an agent who accepts participation in an imperfect pension scheme effectively commits to consuming $\tilde{y}_t = y_t - \alpha_t = \gamma[y_t - a_t^C]$ during the working life and $\tilde{y}_T = y_T + \beta = \gamma[y_T + \sum_{t=1}^{T-1} R^{T-t} a_t^C]$ in retirement. Because the resulting consumption path mimics the shape of the first-best consumption path, the agent has no further incentive to save privately. But because it proportionally lowers the agent's consumption in every period, it may not be welfare improving, relative to the planner's optimum without automatic enrolment.

Notice that for $\gamma = 1$, this scheme collapses to an ideal pension scheme that is necessarily welfare maximising for any agent-type, as discussed above. For a more interesting a case of $\gamma < 1$, the scheme presents the planner with the following trade-off. Should she accept the default pension plan and lower levels of consumption throughout her life cycle, or should she opt out and bear the costs of self-control and planning required to implement her optimal plan? Like in a decision about whether to plan for retirement or not, the planner does a cost-benefit analysis to determine whether she should stay enrolled in the default scheme or not. Intuitively, because the lifetime utility from participating is continuous and strictly increasing in γ, and the lifetime utility from opting out is increasing in ψ and decreasing in ϕ, there should exist a threshold $\underline{\gamma}$ such that an agent characterised by self-control ψ and cognitive costs ϕ accepts a default scheme if and only if $\gamma \geq \underline{\gamma}$. The properties of such threshold for acceptance are summarised in the following result.

Proposition 3.1 (sticky defaults): *For any self-control ψ and cognitive costs ϕ, the planner participates in the imperfect pension scheme if and only if $\gamma \geq \underline{\gamma}$ for some $\underline{\gamma} \in (0, 1]$. Moreover, $\underline{\gamma}$ is (weakly) increasing in ψ and (strictly) decreasing in ϕ.*

In particular, the above proposition implies that for any degree of the scheme's imperfection γ, the default scheme is always more appealing to agents with relatively poor self-control and/or high cognitive costs of planning.

Typically, automatic enrolment is aimed at non-savers. For example, UK law classifies a worker as eligible for automatic enrolment only if they do not already participate in a qualifying workplace pension plan (DWP, 2020). Since rational inaction implies that non-savers are precisely those agents characterised by poor self-control and/or high cognitive costs of planning (Proposition 2.1), the above result suggests a welfare-based explanation for the default option effect - the agents who are most likely to be affected by the change of the default are at the same time least likely to opt out (for a given γ). Patterns consistent with this mechanism are reflected in the 'Attitudes to Pensions' survey. Among those who would be eligible for automatic enrolment, over 70% of respondents had never heard of and knew nothing about the reform. Nevertheless, 68% agreed that it was a good idea and 70% reported that they were likely to stay in the scheme once enrolled (MacLeod et al., 2012).

Furthermore, note that the intuition that agents with high costs of self-control and/or planning are more likely to accept enrolment into an imperfect pension scheme would extend to an alternative specification, where the scheme's imperfection reflects the fact that it induces the agent to save 'too much' (relative to her lifetime resources), rather than proportionally reducing the agent's disposable income in every period. The logic developed above still applies - an agent participates in a default pension scheme if and only if the resulting lifetime utility exceeds her lifetime utility if she were to devise her own plan and bear the cost of exercising self-control. Thus, although automatic enrolment can never induce a 'classical' agent to over-save, this becomes a possibility for credit-constrained agents with pronounced enough behavioural characteristics. This might be particularly relevant in the context of nationwide automatic enrolment in the UK, which imposes age- and income-independent minimum contribution rates and has been shown to have a large impact on participation of young and low-earners.[15]

What if automatic enrolment is introduced at a later stage of the agent's life cycle? Suppose that the agent's state variable in any period $t \geq 1$ corresponds to the current size of the agent's pension pot A_t and her current plan $\hat{C}^t \equiv (\hat{c}_t^t, ..., \hat{c}_{T-1}^t)$, as in section 2.4. Now, opting out is costless and involves implementation of the agent's current plan. If the agent remains enrolled, on the other hand, she needs to bear the cognitive cost of planning in order to deviate from her *new* default consumption path. Then, the result from Proposition 3.1 carries over - holding the agent's income profile and state variables fixed, she is more likely to accept the new default if she has poor self-control and/or high cognitive cost of planning. Of course, agents facing the same lifetime income profile can end up with different pension pots and different plans at t, if they differ in their self-control and costs of planning. We now argue that this only strengthens the main conclusion. To fix ideas, suppose that the classical agent's saving schedule prescribes to save nothing for retirement during the first $\breve{t} < T - 1$ periods and some positive amounts thereafter. Then, if automatic enrolment occurs in this initial stage of the life-cycle, those agents who would start saving at a later date are more likely to opt out than agents who would not, although the two agent-types are initially observationally equivalent. In turn, if automatic enrolments occurs after $\breve{t}$, then the agents who have accumulated some pension wealth by that time are observationally different from the non-savers. It might seem unclear which agent-type should be more likely to accept enrolment into the default scheme now. After all, an agent with positive wealth has lower marginal utility from additional retirement savings and thus may accept participation in the default scheme because it relieves her of bearing the future cost of self-control. However, what drives the planner's decision is the level of lifetime utility associated with opting out, which is always (weakly) greater for an agent with better self-control and lower cognitive costs. Such an agent can always mimic the plan of her low-ψ, high-ϕ counterpart,

[15]For instance, Bourquin, Cribb, and Emmerson (2020) find that automatic enrolment increases participation to around 90%, even among workers with low incomes and in financial difficulty. Andersen and Bhattacharya (2021) study the rationale for mandating young workers to save into illiquid pensions in a model with present-biased preferences and a fixed gap in interest rates on saving and on borrowing.

bearing a lower cost of exercising self-control and saving more, thus attaining higher overall utility. In addition, her resulting consumption in retirement is discretely higher due to already accumulated pension wealth. Then, any adjustments to the agent's plan can only exacerbate the difference in total utility associated with opting out, yielding the following unambiguous prediction.

Corollary 3.1 (mid-life defaults): *For any degree of the default scheme's imperfection γ and the timing of the introduction of automatic enrolment, an agent characterised by better self-control ψ and lower cognitive cost ϕ is more likely to opt out.*

Secondly, in addition to the welfare implications of automatic enrolment, which drive the planner's decision to opt out or accept enrolment, I examine the impact of automatic enrolment on the agent's *total pension wealth*, that is the sum of automatically collected contributions and any other private savings. In the special case of a classical agent, automatic contributions and residual private savings are perfect substitutes, provided that they earn the same return and that the default pension scheme does not induce the agent to save 'too much' (i.e., more than a_t^C). Then, the total pension wealth is invariant to the level of automatic contributions. However, this is no longer true for an agent with a self-control problem and/or positive costs of planning. To isolate the impact of automatic enrolment on wealth accumulation from a welfare-based decision to opt out, define an 'incomplete pension scheme' that is unambiguously welfare-improving, but leaves room for a meaningful decision about additional private savings.

Definition 2: *Let the agent's optimal savings in period $t \leq T-1$ conditional on zero cognitive costs ($\phi = 0$) be denoted by $\bar{a}_t$. An* incomplete pension scheme *collects a contribution of $\alpha_t = \eta \times \bar{a}_t$ in period t and provides an additional benefit of*

$$\beta = \sum_{t=1}^{T-1} R^{T-t}\big[\eta \times \bar{a}_t\big]$$

in period T for some $\eta \in (0,1)$.

It is hopefully clear from the above definition that participation in an incomplete pension scheme always constitutes a welfare-improvement, relative to a non-saving outcome. That is because the scheme is actuarially fair and never leads to over-saving.[16] However, for $\eta < 1$, the agent might still find it worthwhile to save an additional amount privately.

How is the agent's pension wealth accumulation affected, when she is enrolled into an incomplete pension scheme parametrised by η? Intuitively, due the fact that automatic enrolment into

[16] I define an incomplete pension scheme relative to the agent's optimal savings with zero cognitive costs, rather than a_t^C, purely for the conciseness of the following results. Allowing for $\eta > 1$ such that $\eta \times \bar{a}_t \leq a_t^C$ would not change the underlying intuition, but would introduce more cases to consider.

such a scheme strictly increases utility from non-planning, it lowers the incentive to plan and exercise self-control. But whether or not this dis-incentive to plan and save ultimately alters the agent's behaviour, depends on her lifetime utility from active decision-making, which in turn is a function of the agent's self-control and cognitive costs of planning.

To make this point more precise, fix the agent's self-control, and thus $\bar{a}_t$, and define her total wealth accumulated conditional on zero cognitive costs as:

$$\bar{A}_T = \sum_{t=1}^{T-1} R^{T-t}\bar{a}_t$$

The following result states that we can distinguish between three types. First, those with relatively low cognitive costs of planning make an active decision and irrespectively of the automatic enrolment policy end up accumulating pension wealth of $\bar{A}_T$. Second, those with medium cognitive costs do not opt out of the scheme and passively accept the default contribution rate, even though they would save actively in absence of automatic enrolment. This makes them accumulate strictly lower pension wealth of $\beta < \bar{A}_T$. Third, those with high enough cognitive costs are the agents who accumulate no pension wealth in absence of automatic enrolment, but after the change in the default accumulate at least some positive amount $\beta > 0$. Thus, for each of these groups, automatic enrolment has different implications for pension wealth accumulation.

Proposition 3.2 (automatic enrolment and total savings): *Fix the agent's self-control ψ and consider automatic enrolment into an incomplete pension scheme parametrised by η. There exist threshold values for the cognitive cost $\bar{\phi}(\psi) > \underline{\phi}(\psi) \geq 0$, such that*

1. *For $\phi \geq \bar{\phi}(\psi)$, the agent participates in the default scheme and her total pension wealth is strictly above the wealth accumulated in absence of automatic enrolment ("forced saving").*
2. *For $\bar{\phi}(\psi) > \phi \geq \underline{\phi}(\psi)$, the agent participates in the default scheme and her total pension wealth is strictly below the wealth accumulated in absence of automatic enrolment ("discouraged saving").*
3. *For $\underline{\phi}(\psi) > \phi$, the agent opts out of the default scheme and her total pension wealth is unaffected by automatic enrolment.*

Moreover, when $\underline{\phi}(\psi) > 0$, it is strictly decreasing in η.

In the above, $\bar{\phi}(\psi)$ is the cutoff value for cognitive costs, above which an agent left to her own devices is a non-saver. In this case, automatic enrolment into an incomplete pension scheme unambiguously improves welfare and 'forces' the agent to accumulate pension wealth of $\beta > 0$. As discussed at length in section 2, since the ability to exert self-control affects the decision to plan (or not), the cutoff value is written explicitly as a function of ψ. Similarly, $\underline{\phi}(\psi)$ is the cutoff value for cognitive costs, below which an agent finds it optimal to opt out of the default scheme and accumulate wealth of $\bar{A}_T$ privately, as she does in absence of automatic enrolment. For intermediate values of cognitive costs, automatic enrolment generates a 'discouragement effect', whereby an agent who would save $\bar{A}_T$ in absence of automatic enrolment does not find it worthwhile to opt out and ends up with wealth of $\beta < \bar{A}_T$ instead. Intuitively, automatic enrolment reduces the utility gains from saving privately for such 'counterfactually active' savers, resulting in the decision to remain enrolled when $\phi \geq \underline{\phi}(\psi)$. Finally, since the extent to which the default scheme replicates the agent's optimal savings (conditional on zero cognitive costs) reduces her incentive to opt out and re-plan, the threshold for opting out $\underline{\phi}(\psi)$ is monotonically decreasing in η.

Proposition 3.2 highlights the tension at the heart of the policy debate surrounding automatic enrolment. On the one hand, changing the default has a potential to increase pension wealth accumulation of non-savers. On the other hand, it might lower pension wealth of the counterfactually active savers. Naturally, the aggregate effect would depend not only on the extent of the default coverage η, but also on the distribution of types in the population.

One key issue pertaining to the design of an automatic enrolment policy regards the default contribution rate, here captured by parameter η. A commonly voiced concern is that a low default contribution rate might actually *lower* aggregate savings, due to the discouragement effect (e.g., Madrian and Shea, 2001). Does a greater default coverage mechanically lead to greater pension wealth? The answer is: not necessarily. Note that in addition to increasing savings of counterfactual non-savers, an increase in η lowers the cutoff $\underline{\phi}(\psi)$, so that fewer agents are active savers who accumulate $\bar{A}_T$. Again, whether the positive, intensive margin effect dominates the negative, extensive margin effect depends on a distribution of cognitive costs and self-control in the population. This results in the following observation.

Corollary 3.2 (default coverage and total savings): *Consider a population of agents characterised by self-control ψ and some distribution of cognitive costs $\phi \sim [\underline{\phi}, \overline{\phi}]$, such that in absence of automatic enrolment there is a positive mass of non-savers as well as active savers. Then, the aggregate pension wealth accumulated in the population may be increasing, decreasing, or non-monotonic in the default coverage η, depending on the distribution of types.*

Finally, the logic underlying Proposition 3.2 extends to a case where agents are heterogeneous in terms of their self-control and their cognitive costs. Then, the possibility of forced saving at the bottom of the private savings distribution, combined with the possibility of either unaffected or discouraged saving at the top of the distribution produces another important implication of automatic enrolment into an incomplete pension scheme. Namely, such a scheme unambiguously reduces dispersion of the agents' pension wealth.

Corollary 3.3 (automatic enrolment and wealth dispersion): *Consider a population of agents with self-control and cognitive cost parameters distributed over* $\psi \geq \underline{\psi}$ *and* $\phi \in [0, \overline{\phi}]$*, respectively, with positive densities over the entire domain. The agents are otherwise identical. Assume that* $\underline{\psi}$ *and* $\overline{\phi}$ *are such that in absence of automatic enrolment the population consists of non-savers as well as classical agents, i.e.* $A^*_T|_{\psi=\underline{\psi},\phi=\overline{\phi}} = 0$ *and* $A^*_T|_{\psi\to\infty,\phi=0} = b^C$*. Consider the introduction of automatic enrolment into an incomplete pension scheme with a default contribution equal to* $\eta \times a^C_t$ *in period* t *for some* $\eta \in (0, 1]$*. Then, the cross-agent variation in accumulated pension wealth, as measured by the ratio of maximum to minimum total wealth in the population, decreases monotonically in* η *and converges to 1.*

The above corollary illustrates a straightforward mechanism through which automatic enrolment equalises savings outcomes across various agent-types. Intuitively, as the default contribution increases, agents from the bottom of the savings distribution are 'forced' to save more. Agents from the top of the savings distribution, on the other hand, either become 'discouraged' and adhere to the default or opt out and maintain their private savings. The latter is the case when the support of the distribution of ϕ contains 0. Since these properties are maintained for any level of the default coverage η, the variation in pension wealth is monotonically decreasing in η and disappears altogether when $\eta \to 1$, i.e. when the default pension scheme maximises aggregate savings by fully replicating the classical agent's saving schedule.

Note that Corollary 3.3 illustrates a reduction in the dispersion of pension wealth within a given income bracket. Nevertheless, since lower income makes the instance of a corner solution and thus 'forced saving' more likely (for a given ψ and ϕ, see Corollary 2.2), automatic enrolment into an incomplete pension scheme would also reduce the dispersion in average pension wealth across income brackets.

4 Conclusion

This paper analysed the interplay between costly planning and self-control required to follow through with the plan in the context of a life-cycle savings problem. Accounting for the cognitive costs of planning results in the possibility of rational inaction in pension choices. Importantly, the agent's self-control and level of income are identified as inputs into the decision when to formulate a plan, if at all. The model generates realistic predictions in the domain of retirement savings and is used to study how the design of an automatic enrolment policy affects welfare and wealth accumulation. Although this paper focuses on the problem of saving for retirement, the framework developed here could be readily applied to other domains with non-trivial planning and execution stages, e.g. precautionary saving, mortgage repayment, or health-related behaviours.

The scope of the above analysis is limited in a few important ways. First, the level of cognitive costs is assumed to be exogenous and uncorrelated with other individual characteristics. Relaxing this assumption would require more work on the determinants of cognitive strain in financial decisions. Second, I focus on a case of a sophisticated individual, who is perfectly aware of her degree of self-control and cognitive costs. In reality, naiveté about these characteristics may play a role. Moreover, the extent to which the cognitive costs should be regarded as welfare-relevant is generally unclear (Goldin and Reck, 2020). Third, the above model abstracts from an array of factors that have a significant impact on planning for retirement, such as uncertainty, borrowing during the working life, or the timing of retirement. Although recent research does not find evidence that automatic enrolment leads to more indebtedness (Beshears, Choi, Laibson, Madrian, and Skimmyhorn, 2022), a range of factors, such as negative income or health shocks, may result in substantial leakages from retirement saving accounts (Beshears, Choi, Laibson, and Madrian, 2018b). An extension of the model briefly discussed in section 2.4 seems well-suited to studying issues such as optimal re-planning, the interdependence between labour supply and retirement planning, and the trade-off between commitment and flexibility when pension wealth is at least partly illiquid (Amador, Werning, and Angeletos, 2006). Addressing these questions and devising further policy recommendations should be the objectives of future work.

References

Alaoui, L. and A. Penta (2016): "Endogenous Depth of Reasoning," *The Review of Economic Studies*, 83, 1297–1333.

——— (2021): "Cost-Benefit Analysis in Reasoning," *Journal of Political Economy (forthcoming).*

Amador, M., I. Werning, and G.-M. Angeletos (2006): "Commitment vs. Flexibility," *Econometrica*, 74, 365–396.

Ameriks, J., A. Caplin, and J. Leahy (2003): "Wealth Accumulation and the Propensity to Plan," *The Quarterly Journal of Economics*, 118, 1007–1047.

Andersen, T. M. and J. Bhattacharya (2021): "Why mandate young borrowers to contribute to their retirement accounts?" *Economic Theory*, 71, 115–149.

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

Banks, J., C. o'Dea, and Z. Oldfield (2010): "Cognitive Function, Numeracy and Retirement Saving Trajectories," *The Economic Journal*, 120, F381–F410.

Banks, J. and Z. Oldfield (2007): "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving," *Fiscal Studies*, 28, 143–170.

Benartzi, S. and R. Thaler (2007): "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perspectives*, 21, 81–104.

Benhabib, J. and A. Bisin (2005): "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions," *Games and Economic Behavior*, 52, 460–492.

Bernheim, B. D. and D. Taubinsky (2018): "Behavioral Public Economics," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 381–516.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2018a): "Behavioral Household Finance," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 1, 177–276.

——— (2018b): "Potential vs. Realized Savings under Automatic Enrollment," *TIAA Institute Research Dialogue*, 148.

Beshears, J., J. J. Choi, D. Laibson, B. C. Madrian, and W. L. Skimmyhorn (2022): "Borrowing to Save? The Impact of Automatic Enrollment on Debt," *The Journal of Finance*, 77, 403–447.

Beshears, J., K. L. Milkman, and J. Schwartzstein (2016): "Beyond beta-delta: The Emerging Economics of Personal Plans," *American Economic Review*, 106, 430–34.

Bourquin, P., J. Cribb, and C. Emmerson (2020): "Who leaves their pension after being automatically enrolled?" *Institute for Fiscal Studies (IFS) Briefing Note BN272.*

Brocas, I. and J. D. Carrillo (2008): "The Brain as a Hierarchical Organization," *American Economic Review*, 98, 1312–46.

Browning, M. and A. Lusardi (1996): "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, 34, 1797–1855.

Camerer, C. F. (2013): "Goals, Methods, and Progress in Neuroeconomics," *Annual Review of Economics*, 5, 425–455.

Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2002): "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance," *Tax Policy and the Economy*, 16, 67–113.

——— (2003): "Optimal Defaults," *American Economic Review*, 93, 180–185.

Choi, J. J., D. Laibson, B. C. Madrian, A. Metrick, and J. M. Poterba (2004): "For Better or for Worse: Default Effects and 401(k) Savings Behavior," in *Perspectives on the Economics of Aging*, University of Chicago Press.

Choukhmane, T. (2021): "Default Options and Retirement Saving Dynamics," *Working Paper.*

Conlisk, J. (1996): "Why Bounded Rationality?" *Journal of Economic Literature*, 34, 669–700.

Cribb, J. and C. Emmerson (2020): "What happens to workplace pension saving when employers are obliged to enrol employees automatically?" *International Tax and Public Finance*, 27, 664–693.

de la Fuente, A. (2000): *Mathematical Methods and Models for Economists*, Cambridge University Press.

Duckworth, A. L., K. L. Milkman, and D. Laibson (2018): "Beyond Willpower: Strategies for Reducing Failures of Self-Control," *Psychological Science in the Public Interest*, 19, 102–129.

DWP (2020): *Automatic Enrolment evaluation report 2019*, Department for Work and Pensions.

Dynan, K. E., J. Skinner, and S. P. Zeldes (2004): "Do the Rich Save More?" *Journal of Political Economy*, 112, 397–444.

Fudenberg, D. and D. K. Levine (2006): "A Dual-Self Model of Impulse Control," *American Economic Review*, 96, 1449–1476.

——— (2012): "Timing and Self-Control," *Econometrica*, 80, 1–42.

Galperti, S. (2019): "A Theory of Personal Budgeting," *Theoretical Economics*, 14, 173–210.

Goldin, J. and D. Reck (2020): "Optimal Defaults with Normative Ambiguity," *The Review of Economics and Statistics (forthcoming).*

Gul, F. and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69, 1403–1435.

Heath, C., R. P. Larrick, and G. Wu (1999): "Goals as Reference Points," *Cognitive Psychology*, 38, 79–109.

Hurd, M. D. and S. Rohwedder (2013): "Heterogeneity in Spending Change at Retirement," *The Journal of the Economics of Ageing*, 1, 60–71.

John, A. (2020): "When Commitment Fails: Evidence from a Field Experiment," *Management Science*, 66, 503–529.

Keane, M. P. and S. Thorp (2016): "Complex Decision Making: The Roles of Cognitive Limitations, Cognitive Decline, and Aging," in *Handbook of the Economics of Population Aging*, Elsevier, vol. 1, 661–709.

Koch, A. K. and J. Nafziger (2011): "Self-Regulation Through Goal Setting," *Scandinavian Journal of Economics*, 113, 212–227.

——— (2016): "Goals and Bracketing under Mental Accounting," *Journal of Economic Theory*, 162, 305–351.

Kőszegi, B. and F. Matějka (2020): "Choice Simplification: A Theory of Mental Budgeting and Naive Diversification," *The Quarterly Journal of Economics*, 135, 1153–1207.

Lipman, B. L. (1991): "How to Decide How to Decide How to...: Modeling Limited Rationality," *Econometrica*, 1105–1125.

Locke, E. A. and G. P. Latham (1990): *A Theory of Goal Setting & Task Performance.*, Prentice-Hall, Inc.

——— (2002): "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-year Odyssey." *American psychologist*, 57, 705.

Loomes, G. and R. Sugden (1982): "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *The Economic Journal*, 92, 805–824.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2017): "Optimal Financial Knowledge and Wealth Inequality," *Journal of Political Economy*, 125, 431–477.

Lusardi, A. and O. S. Mitchell (2007): "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 54, 205–224.

——— (2011): "Financial literacy and planning: Implications for retirement wellbeing," *NBER Working Paper 17078*.

——— (2014): "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52, 5–44.

MacLeod, P., A. Fitzpatrick, B. Hamlyn, A. Jones, A. Kinver, and L. Page (2012): "Attitudes to Pensions: The 2021 survey," *Department for Work and Pensions Research Report No 813*.

Madrian, B. C. and D. F. Shea (2001): "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 116, 1149–1187.

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

——— (1999b): "Procrastination in Preparing for Retirement," in *Behavioral Dimensions of Retirement Economics*, Brookings Institution Press and Russell Sage Foundation, 125–160.

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

OECD (2019): *Pensions at a Glance 2019*, OECD Publishing, Paris.

——— (2021): *Pensions at a Glance 2021*, OECD Publishing.

Reis, R. (2006): "Inattentive Consumers," *Journal of monetary Economics*, 53, 1761–1800.

Scholz, J. K., A. Seshadri, and S. Khitatrakun (2006): "Are Americans Saving "Optimally" for Retirement?" *Journal of Political Economy*, 114, 607–643.

Shefrin, H. M. and R. H. Thaler (1988): "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry*, 26, 609–643.

Smith, J. P., J. J. McArdle, and R. Willis (2010): "Financial Decision Making and Cognition in a Family Context," *The Economic Journal*, 120, F363–F380.

Sunstein, C. R. and R. H. Thaler (2008): *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press.

Thaler, R. (1985): "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199–214.

Thaler, R. H. (1999): "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12, 183–206.

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

Van Rooij, M. C., A. Lusardi, and R. J. Alessie (2012): "Financial Literacy, Retirement Planning and Household Wealth," *The Economic Journal*, 122, 449–478.

Supplementary Appendix for Online Publication

Appendix A - Alternative specification of psychological disutility

In contrast to the model analysed throughout the paper, consider an alternative specification of the psychological disutility function, where the deviation from the prespecified target is expressed in monetary terms, instead of utils. More specifically:

$$V(c_t, \hat{c}_t) = \begin{cases} \psi\, v\big(c_t - \hat{c}_t\big) & \text{for } c_t \geq \hat{c}_t \\ 0 & \text{for } c_t < \hat{c}_t \end{cases}$$

where $v(\cdot)$ is strictly increasing and strictly convex with $v(0) = 0$ and $v'(0) = 0$, and $\psi > 0$. As before, the agent's instantaneous utility is given by $U(c_t, \hat{c}_t) = u(c_t) - V(c_t, \hat{c}_t)$, and $u(\cdot)$ satisfies the same set of assumptions as in the main text.

Within each period, the doer's choice maximises $U(c_t, \hat{c}_t)$, taking $\hat{c}_t$ as given. Our assumptions imply that the doer's choice, denoted $c_t(\hat{c}_t)$, necessarily lies in the interval $c_t(\hat{c}_t) \in [\hat{c}_t, y_t]$. Interior solutions satisfy the following first-order condition:

$$\psi\, v'\big(c_t - \hat{c}_t\big) \;=\; u'(c_t)$$

Thus, the properties that $c_t(\hat{c}_t)$ is strictly increasing in $\hat{c}_t$ and strictly decreasing in ψ are preserved.

Furthermore, if:

$$\psi\, v'\big(y_t - \hat{c}_t\big) \;<\; u'(y_t)$$

then the no-borrowing constraint is binding at the doer's optimum and $c_t(\hat{c}_t) = y_t$.

The total cost of self-control is given by:

$$\psi\, v\big(c_t(\hat{c}_t) - \hat{c}_t\big)$$

and:

$$\frac{d\,\psi\, v\big(c_t(\hat{c}_t) - \hat{c}_t\big)}{d\,\psi} \;=\; v\big(c_t(\hat{c}_t) - \hat{c}_t\big) \;+\; \psi\, v'\big(c_t(\hat{c}_t) - \hat{c}_t\big)\, \frac{dc_t(\hat{c}_t)}{d\psi} \;<0$$

$$\iff \quad \frac{dc_t(\hat{c}_t)}{d\psi} \;<\; \frac{-v\big(c_t(\hat{c}_t) - \hat{c}_t\big)}{\psi\, v'\big(c_t(\hat{c}_t) - \hat{c}_t\big)}$$

Applying the implicit function theorem, we obtain:

$$\frac{dc_t(\hat{c}_t)}{d\psi} \;=\; \frac{v'\big(c_t - \hat{c}_t\big)}{u''(c_t) - \psi\, v''\big(c_t - \hat{c}_t\big)} < 0$$

and the above inequality becomes:

$$\frac{v'\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)}{u''(c_t(\hat{c}_t))-\psi\, v''\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)} \;<\; \frac{-v\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)}{\psi\, v'\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)}$$

$$\Longleftrightarrow \quad u''(c_t(\hat{c}_t)) \;>\; \frac{\psi\,\Big[v''(c_t(\hat{c}_t)-\hat{c}_t)\times v(c_t(\hat{c}_t)-\hat{c}_t) - (v'(c_t(\hat{c}_t)-\hat{c}_t))^2\Big]}{v(c_t(\hat{c}_t)-\hat{c}_t)}$$

There are two things to notice about this condition. First, it imposes a joint restriction on the curvature of functions $u(\cdot)$ and $v(\cdot)$. Second, because $u(\cdot)$ is assumed to be strictly concave and thus $u''(\cdot) < 0$, the restriction outlined in Assumption 1 is necessary but not sufficient for the above to hold. In that sense, expressing the deviation of the doer's choice from the planner's target in monetary terms and not in utils creates a need to impose a stricter, and arguably more difficult to interpret, functional form assumption in order for the model of self-control to have the desirable property that the total cost of exerting self-control is decreasing in parameter ψ.

Moving on to the planner's problem, the minimum effective target $\underline{\hat{c}}_t$ satisfies:

$$\psi\, v'\big(y_t - \underline{\hat{c}}_t\big) \;=\; u'(y_t)$$

Thus, the agents characterised by poor self-control require a stricter minimum target $\underline{\hat{c}}_t$ to be able to save any positive amount, as before.

Write the planner's problem conditional on a particular planning horizon as:

$$\max_{\hat{c}_\tau,\hat{c}_{\tau+1},\ldots,\hat{c}_{T-1}} \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1}[u(c_t(\hat{c}_t)) - \psi\, v\big(c_t(\hat{c}_t) - \hat{c}_t\big)] + \delta^{T-1}u(A_T)\Big\}, \text{ subject to:}$$

1. $A_T = \sum_{t=\tau}^{T-1} R^{T-t}(y_t - c_t(\hat{c}_t)) + y_T$
2. $\hat{c}_t \leq \underline{\hat{c}}_t$

which can be represented using the following Lagrangian:

$$\mathcal{L}_\tau \;=\; \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1}[u(c_t(\hat{c}_t)) - \psi\, v\big(c_t(\hat{c}_t)-\hat{c}_t\big)] + \delta^{T-1}u\big(\sum_{t=\tau}^{T-1} R^{T-t}(y_t - c_t(\hat{c}_t)) + y_T\big)\Big\} + \sum_{t=\tau}^{T-1} \lambda_\tau[\underline{\hat{c}}_t - \hat{c}_t]$$

For any period $t = \tau, \tau+1, ..., T-1$, the first-order condition associated with $\hat{c}_t$ is:

$$\delta^{t-1}[u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - \psi\, v'\big(c_t(\hat{c}_t) - \hat{c}_t\big)\,[\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - 1] + \delta^{T-1}u'(A_T)R^{T-t}(-\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}) - \lambda_t \;=\; 0$$

$$\Longleftrightarrow \quad u'(c_t(\hat{c}_t))\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - \psi\, v'\big(c_t(\hat{c}_t) - \hat{c}_t\big)\,[\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} - 1] \;=\; (\delta R)^{T-t}u'(A_T)\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} + \tfrac{\lambda_t}{\delta^{t-1}}$$

Taking into account the doer's optimal response to the target, the above simplifies to:

$$u'(c_t(\hat{c}_t)) \;=\; (\delta R)^{T-t} u'(A_T) \tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t} + \tfrac{\lambda_t}{\delta^{t-1}}$$

It follows that in periods when:

$$u'(c_t(\underline{\hat{c}}_t)) \;>\; (\delta R)^{T-t} u'(A_T) \tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}|_{\hat{c}_t=\underline{\hat{c}}_t}$$

it is optimal for the planner to set the optimal lax target $\hat{c}_t = y_t$. Applying the implicit function theorem to this version of the problem, we get:

$$\tfrac{dc_t(\hat{c}_t)}{d\hat{c}_t} \;=\; \frac{-\psi\, v''\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)}{u''(c_t(\hat{c}_t))-\psi\, v''\Big(c_t(\hat{c}_t)-\hat{c}_t\Big)} \;\in (0,1)$$

Since $c_t(\underline{\hat{c}}_t) = y_t$, this condition is different from the one derived in the main text in that the RHS of the inequality is multiplied by a term that is strictly smaller than 1. Moreover, because $\underline{\hat{c}}_t$ is increasing in ψ:

$$\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}|_{\hat{c}_t=\underline{\hat{c}}_t} \;=\; \frac{-\psi\, v''\Big(y_t-\underline{\hat{c}}_t\Big)}{u''(y_t)-\psi\, v''\Big(y_t-\underline{\hat{c}}_t\Big)}$$

is increasing in the agent's self-control as long as $v'''(\cdot) < 0$, but decreasing in self-control otherwise. The former is better suited to capturing the behaviour of a classical agent who implements any target perfectly. In particular, if $\lim_{x\to 0^+} v''(x) = +\infty$, then:

$$\lim_{\psi\to\infty} \tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}|_{\hat{c}_t=\underline{\hat{c}}_t} \;=\; 1$$

so that a classical agent's targets are translated into additional savings dollar-for-dollar. In this case, the condition for not implementing positive savings in period t boils down to the standard Euler equation with a borrowing constraint, but only for agents with perfect self-control.

Furthermore, in periods when the constraint $\hat{c}_t \leq \underline{\hat{c}}_t$ is slack, the planner's optimal target satisfies:

$$u'(c_t(\hat{c}_t)) \;=\; (\delta R)^{T-t} u'(A_T) \tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}$$

which also includes the additional weighting term $\tfrac{d\,c_t(\hat{c}_t)}{d\,\hat{c}_t}$, relative to the standard Euler equation. Thus, this version of the model permits that savings behaviour within a period with positive savings changes smoothly with the self-control parameter ψ, in contrast to being fully captured by the Euler equation. In particular, if:

$$\tfrac{d^2\,c_t(\hat{c}_t)}{d\,\psi\; d\,\hat{c}_t} \;>\; 0,$$

then agents with better self-control save more in period t, ceteris paribus. We have:

$$\frac{d^2\, c_t(\hat{c}_t)}{d\psi\; d\,\hat{c}_t} \;=\; \frac{1}{[u''(\cdot) - \psi\, v''(\cdot)]^2}\left\{ \underbrace{-v''(\cdot)\, u''(\cdot)}_{>0} + \underbrace{\psi \frac{d\, c_t(\hat{c}_t)}{d\,\psi}}_{<0} [v''(\cdot)\, u'''(\cdot) \;-\; v'''(\cdot)\, u''(\cdot)] \right\}$$

This expression is guaranteed to be positive when $u'''(\cdot) \leq \frac{v'''(\cdot)}{v''(\cdot)} u''(\cdot)$, which again reflects jointly the curvatures of the utility function and the psychological disutility. Given our previous assumption that $v'''(\cdot) < 0$, this property is automatically satisfied when $u'''(\cdot) \leq 0$, but the model is also consistent with prudence ($u'''(\cdot) \geq 0$) within the range of parameters permitted by the above condition.

Since $\lim_{\psi \to \infty} c_t(\hat{c}_t) = \hat{c}_t$, under $\lim_{x \to 0^+} v''(x) = +\infty$ we have:

$$\lim_{\psi \to \infty} \frac{d\, c_t(\hat{c}_t)}{d\, \hat{c}_t} \;=\; 1$$

everywhere. That is, the solution to the planner's problem is given by the standard Euler equation, but only when self-control is perfect.

Appendix B - Saving horizon and optimal wealth accumulation

Consider a simple savings model with $T = 3$ periods and suppose that initially savings are permitted only in period 2 and the no-borrowing constraint is slack. Then, the optimal period-2 savings a_2^* satisfy the following Euler equation:

$$u'(y_2 - a_2^*) \;=\; \delta R\, u'(y_3 + R a_2^*)$$

How is the optimal pension wealth affected, when savings are additionally permitted in period 1? With two periods of saving, the optimal consumption path satisfies the following system of Euler equations:

$$\begin{cases} u'(y_1 - a_1^{**}) \;\geq\; (\delta R)^2\, u'(y_3 + R^2 a_1^{**} + R a_2^{**}) \\ u'(y_2 - a_2^{**}) \;=\; \delta R\, u'(y_3 + R^2 a_1^{**} + R a_2^{**}) \end{cases}$$

There are two possibilities. First, if $u'(y_1) \geq (\delta R)^2\, u'(y_3 + R a_2^*)$, then the no-borrowing constraint is binding in period 1 and the optimal savings in period 2 are unaffected, i.e. $a_2^{**} = a_2^*$.

Otherwise, if $u'(y_1) < (\delta R)^2\, u'(y_3 + R a_2^*)$, then the individual Euler equations imply $a_1^{**} > 0$ and $a_2^{**} < a_2^*$. Notice that the sum of these adjustments can only raise period-3 consumption, otherwise we arrive at a contradiction:

$$R^2 a_1^{**} + R a_2^{**} \;<\; R a_2^* \quad \Longrightarrow$$

$$u'(y_2 - a_2^{**}) \;=\; \delta R\, u'(y_3 + R^2 a_1^{**} + R a_2^{**}) \;\;\Longleftrightarrow\;\; a_2^{**} > a_2^*$$

Thus, when an additional period of saving is introduced, the optimal consumption in retirement can only increase. This argument extends iteratively to an arbitrary number of additional periods of saving.

Proof of Lemma 2.1

For an exogenous planning horizon $\tau \neq \varnothing$, the planner's problem is to choose optimal savings targets for each period $t = \tau, \tau+1, ..., T-1$, taking into account the doer's response to the target. As argued in the main text, the conditions (4) and (5) imply that when restricting attention to the domain of effective targets, the agent's degree of self-control has no impact on locally optimal savings (or lack thereof).

However, in order to find a global optimum of the planner's objective function, one must additionally compare the lifetime utility resulting from setting the locally optimal targets to lifetime utilities resulting from setting optimal lax targets $\hat{c}_t = y_t$ in all subsets of periods (and choosing new locally optimal targets in other periods). The reason for this is the observation that the doer's policy function is not differentiable at $\hat{c}_t = \underline{\hat{c}}_t$, where it switches from being determined by the first-order condition (1) to reflecting the binding no-borrowing constraint. Thus, the Lagrangian studied in the main text can only yield the local optimum.

In this step, the agent's self-control starts to play a role. Keep in mind that in any of these local optima, the agent's optimal savings are independent of her self-control ψ. Thus, the pure consumption utility is also independent of ψ. However, by the first-order condition (1), the agents characterised by better self-control (high ψ) implement the same savings by setting themselves targets which are not as strict ($\hat{c}_t$ is increasing in ψ). Then, by Assumption 1, these agents bear a strictly lower cost of self-control associated with positive savings. Therefore, when comparing the lifetime utilities in these local optima, the better the agent's self-control, the less costly it is for her to save positive amounts over a greater number of periods. This implies that at the planner's global optimum, agents with poor self-control set themselves effective targets in (weakly) fewer periods.

To complete the argument note that whenever the standard Euler equation applies, an agent can only accumulate more wealth when saving for retirement over a greater number of working-life periods (Appendix B). Conversely, when saving in a smaller number of periods, the Euler equation implies that an agent saves more per period, but not enough to overturn the relationship between the saving horizon and wealth accumulation.

Proof of Lemma 2.2

Consider the planner's choice between a longer planning horizon τ and a shorter planning horizon $\tau + s$, where planning and saving are postponed by an arbitrary number of periods $s > 0$. Then, the planner optimally postpones planning to $\tau + s$ if and only if:

$$\Upsilon(\tau + s) \;\geq\; \Upsilon(\tau) \quad \Longleftrightarrow$$

$$(1-\delta^s)\phi \ \geq \ \Big\{ \max_{\hat{c}_\tau, \hat{c}_{\tau+1}, \ldots, \hat{c}_{T-1}} \big\{ \sum_{t=\tau}^{T-1} \delta^{t-\tau} [\max_{c_t \leq y_t} u(c_t) - V(c_t, \hat{c}_t)] + \delta^{T-\tau} u(A_T) \big\} \Big\} -$$

$$- \Big\{ \sum_{t=\tau}^{\tau+s-1} \delta^{t-\tau} u(y_t) - \max_{\hat{c}_{\tau+s}, \ldots, \hat{c}_{T-1}} \big\{ \sum_{t=\tau+s}^{T-1} \delta^{t-\tau} [\max_{c_t \leq y_t} u(c_t) - V(c_t, \hat{c}_t)] + \delta^{T-\tau} u(A_T) \big\} \Big\} \quad (6)$$

Notice the following about this condition. First, the RHS of (6) is always non-negative as when optimising over a longer horizon the planner can always set $\hat{c}_t = y_t$ for $t = \tau, ..., \tau + s - 1$, thus at least replicating the level of lifetime utility attained when planning over a shorter horizon. Then, any adjustments to these targets can only yield a utility improvement.

Second, by Lemma 2.1, the lifetime utility associated with each specific planning horizon is (weakly) increasing in self-control ψ. Note that by itself, this does not imply that the utility *improvement* associated with planning over a longer horizon is also increasing in ψ. For example, consider two otherwise identical agents who differ in their self-control. In general, it is possible to find parameter constellations such that an agent with better self-control (high ψ) saves moderate amounts in every period when she starts planning at $\tau + s$, but an agent with poorer self-control (low ψ) does not save at all over the same horizon. But when planning takes place earlier on at τ, a low-ψ agent may eventually find it worthwhile to save a positive amount in one of the periods between τ and $\tau + s - 1$ (for instance, a period with a one-off high disposable income). Thus, her utility improvement might be greater, if the marginal utility improvement from more periods of saving is modest for the high-ψ agent.

Here, Assumption 2 matters. In particular, when the incentives to save for retirement are monotonically increasing over time, situations such as the one outlined above are ruled out. In other words, if a low-ψ agent is not saving over any of the last $T-\tau-s$ periods prior to retirement, she is also not going to save when earlier periods are included in her potential planning horizon. Then, the RHS of (6) is guaranteed to be monotonically increasing in ψ.

To make this point more precise, depending on the agent's self-control and the planning horizon, denote the number of periods in which the planner optimally induces positive savings by $\nu(\psi, \tau) \in \{0, 1, .., T - \tau\}$. The key step is to argue that for $\psi'' > \psi'$, the following holds:

$$\nu(\psi'', \tau) - \nu(\psi'', \tau + s) \ \geq \ \nu(\psi', \tau) - \nu(\psi', \tau + s) \quad (7)$$

Whenever $\nu(\psi', \tau) = \nu(\psi', \tau + s)$, the above is trivially satisfied. Thus, suppose that we have $\nu(\psi', \tau) > \nu(\psi', \tau + s)$. Then, Assumption 2 implies:

$$\nu(\psi', \tau) > \nu(\psi', \tau + s) \ \implies \ \nu(\psi', \tau + s) = T - \tau - s$$

In other words, since the incentives to save for retirement are monotonically decreasing with the distance from retirement, an agent saves a positive amount in a greater number of periods when planning takes place at τ only if she would be saving a positive amount in *every* period over the shorter time horizon (when planning takes place at $\tau + s$).

By the derivation of Lemma 2.1:

$$\nu(\psi', \tau + s) = T - \tau - s \implies \nu(\psi'', \tau + s) = T - \tau - s$$

and then (7) simplifies to $\nu(\psi'', \tau) \geq \nu(\psi', \tau)$, which is always satisfied, also by Lemma 2.1. Thus, we have shown that under Assumption 2, inequality (7) is always satisfied.

Then, to complete the proof of Lemma 2.2, we note that in addition to the RHS of (6) being monotonically increasing in ψ (which makes planning over longer horizons more desirable), the LHS of (6) is strictly increasing in ϕ (which makes planning over longer horizons less desirable). Since these regularities apply to any two arbitrary planning horizons, the cost benefit analysis implies the statement of the Lemma.

As a side note, observe that we could not have invoked the envelope theorem in order to show the above. That is because of lack of regularity in the planner's problem, as discussed in the derivation of Lemma 2.1.

Proof of Proposition 2.1

As in the main text, let $\tau^* \neq \varnothing$ denote the planner's optimal planning horizon, that is:

$$\Upsilon(\tau^*) \equiv$$

$$\max_{\tau \neq \varnothing} \sum_{t=1}^{\tau-1} \delta^{t-1} u(y_t) + \max_{\hat{c}_\tau, \hat{c}_{\tau+1}, \ldots, \hat{c}_{T-1}} \Big\{ \sum_{t=\tau}^{T-1} \delta^{t-1} [u(c_t(\hat{c}_t)) - V(c_t(\hat{c}_t), \hat{c}_t)] + \delta^{T-1} u(A_T) \Big\} - \delta^{\tau-1}\phi$$

where a tie-breaking rule is that when indifferent, the planner chooses a planning horizon with a lower present value of the cognitive costs. Then, $\Upsilon(\tau^*)$ is continuous and strictly decreasing in ϕ. On the other hand, the lifetime utility from non-planning:

$$\Upsilon(\varnothing) \equiv \sum_{t=1}^{T-1} \delta^{t-1} u(y_t) + \delta^{T-1} u(y_T)$$

does not include ϕ, of course. Thanks to continuity and strict monotonicity, we can define $\overline{\phi}(\psi)$ implicitly by:

$$\Upsilon(\tau^*)|_{\phi=\overline{\phi}} = \Upsilon(\varnothing)$$

Under a very mild condition:

$$\Upsilon(\tau^*)|_{\phi=0} > \Upsilon(\varnothing)$$

$\overline{\phi}(\psi)$ is strictly positive and uniquely defined. Then, for any realisation of the cognitive costs $\phi \geq \overline{\phi}$, no planning for retirement takes place at the planner's optimum.

It remains to be shown that $\overline{\phi}(\psi)$ is indeed an increasing function of ψ. This follows from an observation that the lifetime utility attained over any planning horizon $\Upsilon(\tau)$, $\tau \neq \varnothing$, is weakly increasing in ψ (Lemma 2.1). Then, $\Upsilon(\tau^*)$ is also weakly increasing in ψ.

As a final step, note that this implies that independent of the planning horizon chosen at the planner's optimum, the marginal level of cognitive costs that make an agent a non-planner must be strictly increasing in self-control ψ. By contradiction, consider two levels of self-control $\psi'' > \psi'$ and suppose that $\overline{\phi}(\psi') > \overline{\phi}(\psi'')$. Then, the definition of $\overline{\phi}$ yields:

$$\Upsilon(\tau^*)|_{\psi=\psi',\phi=\overline{\phi}(\psi')} \;=\; \Upsilon(\tau^*)|_{\psi=\psi'',\phi=\overline{\phi}(\psi'')}$$

which is a contradiction, because $\Upsilon(\tau^*)$ is weakly increasing in ψ and strictly decreasing in ϕ. Thus we must have $\overline{\phi}(\psi'') \geq \overline{\phi}(\psi'')$. Moreover, $\overline{\phi}(\psi)$ is strictly increasing whenever at least one target is effective at the planner's optimal planning horizon, which is also a rather weak condition.

Proof of Corollary 2.1

Conditional on the cognitive cost ϕ, agents characterised by better self-control ψ induce positive savings in a greater number of periods for any planning horizon (Lemma 2.1) and are less likely to not plan at all (Proposition 2.1), which necessarily results in higher consumption in retirement (Appendix B). Moreover, if the conditions outlined in Assumption 2 are satisfied, better self-control allows an agent to plan for retirement earlier on (Lemma 2.2), again resulting in higher pension wealth accumulation.

Even if Assumption 2 is not met, and thus it is conceivable that agents with lower ψ start saving in earlier periods, they must necessarily be under-consuming in retirement. Otherwise, an agent with better self-control could achieve at least the same degree of consumption smoothing at a lower cost of self-control by mimicking the behaviour of a low-ψ agent, contradicting the optimality of her plan.

Conditional on the degree of self-control ψ, an increase in cognitive costs ϕ makes an agent plan over shorter horizons (derivation of Lemma 2.2), which can only lower pension wealth accumulation (Appendix B).

Proof of Corollary 2.2

When the utility function is CRRA, for some $\theta > 0$ we have:

$$u(x) = \tfrac{x^{(1-\theta)}-1}{1-\theta} \quad \Longrightarrow \quad u'(x) = x^{-\theta}$$

and thus the optimality conditions in a standard consumption-savings problem are unaffected by scaling up the entire lifetime income profile by a constant factor μ, in the sense that the same optimal savings rate applies in every period, see Euler equations (4) and (5).

We now show that the RHS of (6) is strictly increasing in the scale of lifetime income profile μ. This would imply that as μ increases, an agent becomes more likely to adopt a longer planning horizon, which by the result in Appendix B can only lead to higher wealth accumulation (holding the income profile constant). Then, whenever the planning horizon is extended, the wealth-to-income ratio attained at the planner's optimum is (weakly) increasing.

It should be clear that the difference in pure consumption utilities captured in the RHS of (6) is inflated when μ increases. What about the cost of exercising self-control $V(c_t, \hat{c}_t)$? Given that the Euler equations yield the same solution for the optimal savings rates, the relevant comparison is between the cost of implementing consumption level c_t when the income is y_t and the cost of implementing consumption level μc_t when the income is μy_t. Let the target required to induce the doer to consume c_t be denoted by $\hat{c}_t(c_t)$. Then:

$$V(\mu c_t, \hat{c}_t(\mu c_t)) \; = \; \psi v(u(\mu c_t) - u(\hat{c}_t(\mu c_t)))$$

$$\implies \quad \frac{dV(\mu c_t, \hat{c}_t(\mu c_t))}{d\mu} \; = \; \psi v'(u(\mu c_t) - u(\hat{c}_t(\mu c_t)))\left[u'(\mu c_t)c_t - u'(\hat{c}_t(\mu c_t))\frac{d\hat{c}_t(\mu c_t)}{d\mu c_t}c_t\right]$$

where from the doer's optimal response (1), we have $\psi v'(u(\mu c_t) - u(\hat{c}_t(\mu c_t))) = 1$. What is more, the implicit function theorem applied in the main text gives:

$$\frac{d\hat{c}_t(\mu c_t)}{d\mu c_t} \; = \; \frac{u'(\mu c_t)}{u'(\hat{c}_t(\mu c_t))}$$

which results in:

$$\frac{dV(\mu c_t, \hat{c}_t(\mu c_t))}{d\mu} \; = \; 0$$

Thus, the absolute cost of self-control associated with implementing the desired consumption plan is independent of the scaling parameter μ. In sum, this implies that the RHS of (6) is indeed strictly increasing in μ and the statement of the corollary follows.

Proof of Corollary 2.3

First note that according to the standard Euler equations (4) and (5), a small increase in y_T results in a small downward adjustment in savings levels across periods with positive savings. In other words, the agent's savings change continuously with y_T, up to the point when the no-borrowing constraint starts binding.

For an agent with either positive cost of planning or imperfect self-control, this implies that an increase in y_T strictly reduces the RHS of (6). That is because, irrespective of the degree of self-control ψ, the first-best levels of savings and thus the marginal benefit from smoothing consumption over a longer horizon become reduced, while the cost of self-control associated with inducing positive savings in any period is unaffected. As a result, an increase in y_T might make an agent save over a shorter horizon, resulting in strictly lower pension wealth (Appendix B).

In addition, when Assumption 2 is met, the proof of Lemma 2.2 implies that for any given planning horizon, the agents with poor self-control are most likely to shorten their planning horizon following an increase in y_T. That is because in this case, the RHS of (6) is increasing in ψ, in addition to being strictly decreasing in y_T.

Proof of Proposition 3.1

As discussed in the main text, when the agent accepts participation in the imperfect pension scheme, she attains the lifetime utility of:

$$\Upsilon^D|_\gamma \;\equiv\; \sum_{t=1}^{T-1} \delta^{t-1} u(\gamma(y_t - a_t^C)) + \delta^{T-1} u(\gamma(y_t + \sum_{t=1}^{T-1} R^{T-t} a_t^C))$$

which is strictly increasing in γ, but independent of self-control ψ and cognitive costs ϕ. Moreover, $\Upsilon^D|_\gamma$ attains the first-best outcome for $\gamma = 1$.

We already know from the derivation of Proposition 2.1 that the lifetime utility attained by the planner in a SR-perfect NE, denoted $\Upsilon(\tau^*)$, is weakly increasing in ψ and strictly decreasing in ϕ. Then, the planner accepts participation in the imperfect scheme if and only if:

$$\Upsilon^D|_\gamma \;\geq\; \Upsilon(\tau^*)$$

and the statement of the proposition follows. Note also that since we have

$$\Upsilon^D|_{\gamma=1} \;\geq\; \Upsilon(\tau^*)$$

and

$$\Upsilon^D|_{\gamma=0} \;<\; \Upsilon(\tau^*)$$

for all ψ, ϕ, for any agent-type the threshold for accepting the default scheme $\underline{\gamma}$ is unique and strictly positive.

Proof of Corollary 3.1

The statement of the corollary follows from the observation that the agent's continuation payoff, in case she opts out of the default scheme, is weakly increasing in self-control ψ and decreasing in cognitive costs ϕ. In addition, accounting for endogeneity of the wealth accumulation path only strengthens that conclusion in the sense that an agent with higher current pension wealth, better self-control, and lower cognitive costs can always attain a discretely higher continuation payoff than her relevant counterpart.

Proof of Proposition 3.2

We define the cutoff for the cognitive costs $\overline{\phi}(\psi)$ above which an agent left to her own devices is a non-saver in the same way as in the proof of Proposition 3.1, i.e.:

$$\Upsilon(\tau^*)|_{\phi=\overline{\phi}} \;=\; \Upsilon(\varnothing)$$

Now, when the agent accepts the participation in an incomplete pension scheme characterised by η, she attains the lifetime utility of:

$$\Upsilon^D|_\eta \;\equiv\; \sum_{t=1}^{T-1} \delta^{t-1} u(y_t - \eta \overline{a}_t) + \delta^{T-1} u(y_t + \sum_{t=1}^{T-1} R^{T-t}(\eta \overline{a}_t))$$

which is strictly increasing in $\eta \in [0,1]$, but independent of self-control ψ and cognitive cost ϕ. Then, the cutoff for the cognitive costs $\underline{\phi}(\psi)$ above which an agent prefers the default to saving $\overline{a}_t$ in each period privately is defined implicitly by:

$$\Upsilon(\tau^*)|_{\phi=\underline{\phi}} \;=\; \Upsilon^D|_\eta$$

Since $\Upsilon^D|_{\eta=0} = \Upsilon(\varnothing)$ and $\Upsilon(\tau^*)$ is strictly decreasing in ϕ, for any $\eta > 0$ we have $\underline{\phi}(\psi) < \overline{\phi}(\psi)$. Furthermore, whenever the following is satisfied for a given $\eta > 0$:

$$\Upsilon(\tau^*)|_{\phi=0} \;>\; \Upsilon^D|_\eta$$

then we have $\underline{\phi}(\psi) > 0$ and the two cutoffs are unique. This condition may fail, however, if the agent's cost of self-control and η are high enough, in which case $\underline{\phi}(\psi) = 0$.

Finally, because $\Upsilon^D|_\eta$ is strictly increasing in η, the above condition implies that $\underline{\phi}(\psi)$ is strictly decreasing in η whenever it takes strictly positive values.

Proof of Corollary 3.2

Let the distribution function $F(\cdot)$ represent the distribution of cognitive-cost types in a population. To focus attention, assume that $F(\cdot)$ is differentiable, although that is not necessary for the corollary to hold true. Then, a mass of agents who are non-savers in absence of automatic enrolment is $1 - F(\overline{\phi})$. For those agents, automatic enrolment increases pension wealth by $\eta \overline{A}_T$.

A mass $F(\overline{\phi}) - F(\underline{\phi})$ of agents are counterfactual active savers who nonetheless accept the default. For each of those agents, automatic enrolment reduces pension wealth by $(1-\eta)\overline{A}_T$.

A remaining mass of agents $F(\underline{\phi})$ are not affected by the change in the default, as they always choose their savings actively. Thus, the total impact of automatic enrolment into an incomplete pension scheme on aggregate wealth accumulation is given by:

$$\Delta \;=\; (1 - F(\overline{\phi})) \times \eta \overline{A}_T \;-\; (F(\overline{\phi}) - F(\underline{\phi})) \times (1-\eta)\overline{A}_T$$

Since $\overline{\phi}$ is independent of η, but $\underline{\phi}$ is strictly decreasing in η, the marginal impact of the default coverage on aggregate wealth is:

$$\frac{d\Delta}{d\eta} = (1 - F(\overline{\phi}))\overline{A}_T + (F(\overline{\phi}) - F(\underline{\phi}))\overline{A}_T + f(\underline{\phi})\frac{d\underline{\phi}}{d\eta}(1-\eta)\overline{A}_T =$$

$$= (1 - F(\underline{\phi}))\overline{A}_T + f(\underline{\phi})\frac{d\underline{\phi}}{d\eta}(1-\eta)\overline{A}_T$$

The first term is strictly positive and captures the fact that a higher coverage of the default scheme increases wealth of passive savers. The second term is strictly negative (as long as $\underline{\phi}(\psi) > 0$)) and captures the fact that a higher default coverage increases the mass of passive savers accepting a strictly lower savings rate than in a counterfactual setting without automatic enrolment.

Then, the statement of the corollary follows from the observation that depending on the parametrisation of the problem, either the positive or the negative effect can be dominant for any η.

Proof of Corollary 3.3

Since the support of the types distribution is assumed to contain classical agents with zero cognitive costs and perfect self-control, the maximum wealth accumulated in this population is equal to b^C, irrespective of η. Moreover, the minimum wealth accumulated in this population is ηb^C corresponding to the group of passive savers. Then, as $\eta \to 1$, the ratio $b^C/\eta b^C$ is of course monotonically decreasing in η and converges to 1.

PREVIOUS DISCUSSION PAPERS

384 Sulka, Tomasz, Planning and Saving for Retirement, March 2022.

383 Cattan, Sarah, Kamhöfer, Daniel A., Karlsson, Martin and Nilsson, Therese, The Long-term Effects of Student Absence: Evidence from Sweden, March 2022.

382 Martin, Simon and Rasch, Alexander, Collusion by Algorithm: The Role of Unobserved Actions, March 2022.

381 Haucap, Justus, Nedic, Radivoje and Şimşek, Talha, An Empirical Analysis of German Casino Locations, March 2022.

380 Haucap, Justus, Heldman, Christina and Rau, Holger A., Gender and Collusion, March 2022.

379 Schain, Jan Philip, Foreign Institutional Investors and the Great Productivity Slowdown, February 2022.

378 Neyer, Ulrike and Stempel, Daniel, How Should Central Banks React to Household Inflation Heterogeneity?, January 2022.

377 Döpper, Hendrik and Rasch, Alexander, Combinable Products, Price Discrimination, and Collusion, January 2022.

376 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, Self-Control and Unhealthy Body Weight: The Role of Impulsivity and Restraint, January 2022.

375 Martin, Simon and Schmal, W. Benedikt, Collusive Compensation Schemes Aided by Algorithms, December 2021.

374 Kouli, Yaman and König, Jörg, Measuring European Economic Integration 1880 - 1913 – A New Approach, December 2021.

373 Hunold, Matthias and Schad, Jannika, Single Monopoly Profits, Vertical Mergers, and Downstream Entry Deterrence, December 2021.

372 Werner, Tobias, Algorithmic and Human Collusion, November 2021.

371 Jovanovic, Dragan, Wey, Christian and Zhang, Mengxi, On the Social Welfare Effects of Runner-up Mergers in Concentrated Markets, September 2021.
Published in: Economics Bulletin, (2021), Article 201.

370 Jeschonneck, Malte, Collusion among Autonomous Pricing Algorithms Utilizing Function Approximation Methods, August 2021.

369 Gösser, Niklas, Gürer, Kaan, Haucap, Justus, Meyring, Bernd, Michailidou, Asimina, Schallbruch, Martin, Seeliger, Daniela and Thorwarth, Susanne, Total Consumer Time – A New Approach to Identifying Digital Gatekeepers, August 2021.

368 Fischer, Kai, Reade, J. James and Schmal, W. Benedikt, The Long Shadow of an Infection: COVID-19 and Performance at Work, August 2021.

367 Suedekum, Jens, Place-Based Policies – How to Do Them and Why, August 2021.

366 Heiss, Florian, Ornaghi, Carmine and Tonin, Mirco, Inattention vs Switching Costs: An Analysis of Consumers' Inaction in Choosing a Water Tariff, July 2021.

365 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, Sophistication about Self-Control, July 2021.

364 Bie, Xiaodong and Ciani, Andrea, Born Similar, Develop Apart: Evidence on Chinese Hybrid Exporters, July 2021.

363 Ali, Nesma and Stiebale, Joel, Foreign Direct Investment, Prices and Efficiency: Evidence from India, July 2021.

362 Banerjee, Ritwik, Ibanez, Marcela, Riener, Gerhard and Sahoo, Soham, Affirmative Action and Application Strategies: Evidence from Field Experiments in Columbia, April 2021.

361 Wellmann, Nicolas and Czarnowske, Daniel, What Would Households Pay for a Reduction of Automobile Traffic? Evidence From Nine German Cities, March 2021.

360 Haucap, Justus, Moshgbar, Nima and Schmal, Wolfgang Benedikt, The Impact of the German "DEAL" on Competition in the Academic Publishing Market, March 2021.
Published in: Managerial and Decision Economics, 42 (2021), pp. 2027-2049.

359 Korff, Alex, Competition in the Fast Lane – The Price Structure of Homogeneous Retail Gasoline Stations, January 2021.

358 Kiessling, Lukas, Chowdhury, Shyamal, Schildberg-Hörisch, Hannah and Sutter, Matthias, Parental Paternalism and Patience, January 2021.

357 Kellner, Christian, Le Quement, Mark T. and Riener, Gerhard, Reacting to Ambiguous Messages: An Experimental Analysis, December 2020.

356 Petrishcheva,Vasilisa, Riener, Gerhard and Schildberg-Hörisch, Hannah, Loss Aversion in Social Image Concerns, November 2020.

355 Garcia-Vega, Maria, Kneller, Richard and Stiebale, Joel, Labor Market Reform and Innovation: Evidence from Spain, November 2020.
Published in: Research Policy, 50 (2021), 104213.

354 Steffen, Nico, Economic Preferences, Trade and Institutions, November 2020.

353 Pennerstorfer, Dieter, Schindler, Nora, Weiss, Christoph and Yontcheva, Biliana, Income Inequality and Product Variety: Empirical Evidence, October 2020.

352 Gupta, Apoorva, R&D and Firm Resilience During Bad Times, October 2020.

351 Shekhar, Shiva and Thomes, Tim Paul, Passive Backward Acquisitions and Downstream Collusion, October 2020.
Forthcoming in: Economics Letters.

350 Martin, Simon, Market Transparency and Consumer Search – Evidence from the German Retail Gasoline Market, September 2020.

349 Fischer, Kai and Haucap, Justus, Betting Market Efficiency in the Presence of Unfamiliar Shocks: The Case of Ghost Games during the COVID-19 Pandemic, August 2020.
Substantially revised version forthcoming in: Kyklos.

348 Bernhardt, Lea, Dewenter, Ralf and Thomas, Tobias, Watchdog or Loyal Servant? Political Media Bias in US Newscasts, August 2020.

347 Stiebale, Joel, Suedekum, Jens and Woessner, Nicole, Robots and the Rise of European Superstar Firms, July 2020.

346 Horst, Maximilian, Neyer, Ulrike and Stempel, Daniel, Asymmetric Macroeconomic Effects of QE-Induced Increases in Excess Reserves in a Monetary Union, July 2020.

345 Riener, Gerhard, Schneider, Sebastian O. and Wagner, Valentin, Addressing Validity and Generalizability Concerns in Field Experiments, July 2020.

344 Fischer, Kai and Haucap, Justus, Does Crowd Support Drive the Home Advantage in Professional Soccer? Evidence from German Ghost Games during the COVID-19 Pandemic, July 2020.
Published in: Journal of Sports Economics, 22 (2021), pp. 982-1008.
.

343 Gösser, Niklas and Moshgbar, Nima, Smoothing Time Fixed Effects, July 2020.

342 Breitkopf, Laura, Chowdhury, Shyamal, Priyam, Shambhavi, Schildberg-Hörisch, Hannah and Sutter, Matthias, Do Economic Preferences of Children Predict Behavior?, June 2020.

341 Westphal, Matthias, Kamhöfer, Daniel A. and Schmitz, Hendrik, Marginal College Wage Premiums under Selection into Employment, June 2020.

340 Gibbon, Alexandra J. and Schain, Jan Philip, Rising Markups, Common Ownership, and Technological Capacities, April 2021 (First Version June 2020).

339 Falk, Armin, Kosse, Fabian, Schildberg-Hörisch, Hannah and Zimmermann, Florian, Self-Assessment: The Role of the Social Environment, May 2020.

338 Schildberg-Hörisch, Hannah, Trieu, Chi and Willrodt, Jana, Perceived Fairness and Consequences of Affirmative Action Policies, April 2020.

337 Avdic, Daniel, de New, Sonja C. and Kamhöfer, Daniel A., Economic Downturns and Mental Wellbeing, April 2020.
Published in: European Economic Review, 140 (2021), 103915 under the title "Economic Downturns and Mental Health in Germany".

336 Dertwinkel-Kalt, Markus and Wey, Christian, Third-Degree Price Discrimination in Oligopoly When Markets Are Covered, April 2020.
Forthcoming in: Journal of Industrial Economics.

335 Dertwinkel-Kalt, Markus and Köster, Mats, Attention to Online Sales: The Role of Brand Image Concerns, April 2020.
Published in: Journal of Economics and Management Strategy, 31 (2022), pp. 64-89.

334 Fourberg, Niklas and Korff, Alex, Fiber vs. Vectoring: Limiting Technology Choices in Broadband Expansion, April 2020.
Published in: Telecommunications Policy, 44 (2020), 102002.

333 Dertwinkel-Kalt, Markus, Köster, Mats and Sutter, Matthias, To Buy or Not to Buy? Price Salience in an Online Shopping Field Experiment, April 2020.
Revised version published in: European Economic Review, 130 (2020), 103593.

332 Fischer, Christian, Optimal Payment Contracts in Trade Relationships, February 2020.

331 Becker, Raphael N. and Henkel, Marcel, The Role of Key Regions in Spatial Development, February 2020.

330 Rösner, Anja, Haucap, Justus and Heimeshoff, Ulrich, The Impact of Consumer Protection in the Digital Age: Evidence from the European Union, January 2020.
Published in: International Journal of Industrial Organization, 73 (2020), 102585.

Older discussion papers can be found online at:
http://ideas.repec.org/s/zbw/dicedp.html

Heinrich-Heine-Universität Düsseldorf

Düsseldorfer Institut für Wettbewerbsökonomie (DICE)

Universitätsstraße 1, 40225 Düsseldorf

ISSN 2190-992X (online)
ISBN 978-3-86304-383-4

www.dice.hhu.de