Modelling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions

Jess Benhabib   Alberto Bisin
NYU   NYU

First Draft: July 2002; This Draft: February 2004[*]

Abstract

We provide a new model of consumption-saving decisions which explicitly allows for internal commitment mechanisms and self-control. Agents have the ability to invoke either automatic processes that are susceptible to the temptation of 'over-consuming,' or alternative control processes which require internal commitment but are immune to such temptations. Standard models in behavioral economics ignore such internal commitment mechanisms. We justify our model by showing that much of its construction is consistent with dynamic choice and cognitive control as they are understood in cognitive neuroscience.

The dynamic consumption-saving behavior of an agent in the model is characterized by a simple consumption-saving goal, a propensity to consume out of wealth which is independent of any realized temptation, and a cut-off rule for invoking control processes to inhibit automatic processes and implement the consumption-saving goal. We compare the behavior induced by our model with that induced by standard behavioral models where agents have no internal commitment ability. We discuss empirical tests of our model with available individual consumption data and we suggest critical tests with brain-imaging and experimental data.

1 Introduction

Consider the standard economic approach to the study of consumption and saving behavior, after Friedman's Permanent Income Hypothesis and Modigliani's Life-Cycle Hypothesis (Friedman (1957) and Modigliani-Brumberg (1954), respectively). It involves an agent choosing a feasible consumption plan c_t to maximize his present *exponentially* discounted utility. Recently, behavioral economists have criticized this approach[1] on the basis of a vast amount of empirical evidence in experimental psychology indicating that agents may have a preference for present consumption that cannot be rationalized with exponential discounting.[2] They have suggested an alternative specification of discounting, *quasi-hyperbolic discounting,*[3] which rationalizes the preference for present consumption as a form of time inconsistency.[4]

When preferences are time inconsistent, agents' decisions are not only determined by rationality: At each stage agents must make decisions based on expectations regarding their own future decisions, which will be based on different preference orderings than the present one. Such expectations must therefore be determined in equilibrium. The behavioral economics literature models dynamic decisions as a sequential game between different 'selves', each one choosing at a different time, and it restricts the analysis to Markov Perfect Nash equilibria.[5] By considering only Markovian strategies of a game between present and future selves the behavioral economics literature implicitly models agents as lacking any form of internal psychological commitment ability, or self-control.[6,7]

This is hardly justified. First of all the experimental evidence which contradicts exponential discounting does not automatically deliver an alternative theory of dynamic choice: these experiments are explicitly designed to avoid choices that require commitment or self-control.[8] Moreover, a vast theoretical and experimental literature in psy-

---

[1]See e.g., Laibson (1996), O'Donoghue-Rabin (1999).

[2]See, e.g., Ainsle (1992), (2001), Ainsle-Haslam (1992), Frederick-Loewenstein-O'Donoghue (2002) for comprehensive surveys.

[3]Psychologists favor a related specification, *hyperbolic discounting;* see e.g., Herrnstein (1961), de Villiers-Herrnstein (1976), and Ainsle (1992).

[4]Of course, quasi-hyperbolic discounting (or even, more generally, time inconsistency) is not the only possible way to rationalize the experimental evidence. Rubinstein (2000) shows how such evidence is consistent with a specific form of procedural rationality, and Gul-Pesendorfer (2001) rationalize it with preferences over sets of actions, under standard rationality axioms; see Gul-Pesendorfer (2000).

[5]See the special issue of the *Journal of Economic Perspectives*, 2001, on the topic, and the references therein.

[6]We use internal commitment and self-control essentially as synonymous in this paper, following the standard use in economics and psychology. This is not to imply that internal commitment mechanisms are governed by a 'self.' In fact the cognitive control models we adopt as foundations of our analysis are careful in not requiring a 'self' or a 'homunculus;' see the introduction to Monsell-Driver (2000).

[7]But see Benabou-Tirole (2000), which exploits information asymmetries across different selves, and Bayesian inference methods in the strategic interaction between the selves, to develop a theory of self-control.

[8]The design of these experiments aims to 'uncover natural spontaneous preferences' (Ainsle, (2001),

chology does in fact study the problem of dynamic choice, and identifies various internal commitment and self-control strategies that agents use to implement their objectives.[9] It is our contention therefore that the dynamic choices of agents with time inconsistent preferences cannot be properly understood without an explicit analysis of the dynamic commitment strategies involving self-control.

In this paper we provide a new model of consumption-saving decisions which explicitly allows for internal commitment mechanisms and self-control. We justify our model by showing that much of its construction is consistent with cognitive control as it is understood in cognitive neuroscience.[10] Agents have the ability to either invoke *automatic processes* that are susceptible to impulses or temptations, or alternative *control processes* which are immune to such temptations. Controlled process in our model induce the agent to implement a set of goals, determined independently of impulses or temptations associated with the specific choice problem. The differential activation of the automatic and controlled processing determines which process is responsible for the agent's choice. It depends on the learned future expected rewards to the actions induced by the two processes. The neurobiological foundation of the basic postulate of this analysis, that internal commitment and self-control in dynamic choice operate as a specific form of cognitive control, has never been tested with imaging data. We identify a critical dynamic choice experiment that can generate reaction time and brain imaging data to directly test this postulate.

Based on this model of internal commitment and self-control, we develop a theory of dynamic decision-making which we apply to a standard consumption-saving problem. Agents trade-off 'excessive' and 'impulsive' immediate consumption with a consumption-saving rule requiring the exercise of self-control for its implementation. In particular, the present bias in the model derives from stochastic temptations that affect the agents' consumption-saving choice each period. Self-control requires actively maintaining attention to a specific goal, e.g., an optimal consumption-saving rule that is unaffected by temptations. Such a consumption-saving rule, to be implemented, requires inhibitory connections that become stronger the higher is the cognizance of expected regret in response to 'impulsive' and immediate consumption.

The behavior of an agent facing conflicting preference representations over his consumption-saving choice can be simply summarized. At times the agent allows temptations to affect his consumption-saving behavior by letting the automatic choice prevail, if this choice does not perturb his underlying consumption-saving plan too much, and does not have large permanent effects on his prescribed wealth accumulation pattern. When evaluating

---

pg. 33), that is, to 'observe situations where the subject is not challenged to exercise self-control' (Ainsle (1992), pg. 70).

[9]See e.g., Kuhl-Beckmann (1985) for a survey, and Gollwitzer-Bargh (1996) for a collection of essays on the topic.

[10]See Miller-Cohen (2001) and O'Reilly-Munakata (2000) for comprehensive surveys of the literature on cognitive control.

the effects of a deviation from prescribed consumption-saving patterns to accomodate a temptation, agents do anticipate that such a temptation will in fact be followed by other ones in the future, and their consumption-saving rule will reflect this anticipation.

We derive some implications of our cognitive model of self-control to better understand how changes in the external environment affect consumption-saving behavior. For example, we show that an environment with larger temptations is characterized by a higher probability that self-control is exercised and temptations are inhibited. On the other hand, in such an environment, agents set less ambitious saving goals, that is they consume a larger fraction of their accumulated wealth each time self-control is exercised. We show that an agent with lower cognitive control abilities, or, equivalently, an agent whose attention is consumed by other important cognitive tasks, exercises self-control less frequently, and furthermore, sets less ambitious goals in attempting to inhibit temptations. We study the *complexity* of the consumption-saving goal that agents set for themselves. Psychologists constantly remark that the 'complexity' of goals reduces agents' effectiveness in tasks of self-regulation and in particular in tasks of self-control.[11] According to this view, a cognitive task is simpler to implement the simpler are the goals, e.g., because simple goals do not require exclusive attention. In such an environment, we characterize conditions under which an agent would gain from setting a simpler consumption-saving goal, e.g., a constant saving rule, as opposed to a 'complex' goal, that is one contingent on the rate of return on savings. We show that the simple consumption-saving goal may be preferred to the complex goal. More interestingly, the simple goal tends to be preferred if the rate of return is small enough, as in this case self-control is of little use, and it is a dominant choice for the agent to consume a large fraction of his wealth each period. The simpler goal will also tend to be preferred, for instance, if temptations grow large on average. This is because when temptations are large enough both the complex and the simple goal will optimally induce inhibition of the automatic processing most of the time, but the simpler goal is easier to actively maintain.

Finally, we compare the consumption-saving behavior implied by our model with that implied by standard behavioral models where agents have no internal commitment ability. In section 3.4 we identify critical empirical tests of our model against these alternatives with data on individual consumption, portfolio composition, and asset prices. We survey the existing evidence and document the following: 'excess sensitivity' of consumption is greatly reduced in the case of large windfall gains, liquid assets are traded at a relatively large premium, agents tend not to consume nor borrow out of their real-estate equity, nor out of future life insurance benefits. We argue that this evidence in fact supports our cognitive consumption-saving model.

---

[11]See for instance Baumeister-Heatherton-Tice (1994), Gollwitzer-Bargh (1996), and Kuhl-Beckmann (1985).

2   A cognitive model of dynamic choice and control

In this section we introduce the notion of cognitive control and outline the theoretical and empirical literature in the cognitive sciences that will form the foundation of our analysis of dynamic choice. We rely on models of cognitive control in neuroscience which aim at developing a general integrated theory of cognitive behavior based on the function of the prefrontal cortex, as Braver-Cohen-Servan Schreiber (1995); see also Miller-Cohen (2001) and O'Reilly-Munakata (2000) for surveys. The core of such models is the classical distinction between *automatic* and *controlled* processing, as articulated e.g., in Shiffrin-Schneider (1977), Norman-Shallice (1980), Shallice (1988). Automatic processes are based on the learned association of a specific response to a collection of cues, and underlie classical conditioning and Pavlovian responses.[12] Controlled processes are instead based on the activation, maintenance, and updating of active goal-like representations in order to influence cognitive procedures, and possibly to inhibit automatic responses.[13] Cognitive control is the result of differential activations of automatic and controlled processing pathways. An *executive function,* or *supervisory attention system*, modulates the activation levels of the different processing pathways, based on the learned representation of expected future rewards.[14] Cognitive control might fail, as controlled processes fail to inhibit automatic reactions, because actively maintaining the representation of a goal is costly, due to the severe biological limitations of the activation capacity of the supervisory attention system of the cortex.[15,16]

---

[12]Automatic processes are associated to the activation of various areas of the posterior cortex; see e.g., Schultz-Dayan-Montague (1997).

[13]Controlled processes are associated to sustained neural activity in the pre-frontal cortex during cognitive tasks; see Cohen-Perlstein-Braver-Nystrom-Noll (1997) and Prabhakaran-Narayanan-Zhao-Gabrieli (2000).

[14]The areas of the brain specialized in representing and predicting future rewards are the midbrain nuclei the ventral tegmental area (VTA) and the substantia nigra; see Schultz-Apicella-Romo-Scarnati (1995) for neural recording studies, Bechara-Tranel-Damasio-Damasio (1996) for clinical studied of patients with brain lesions, and Schultz (1998) for a survey. The biological processes which constitute the supervisory attention system modulating the activation of automatic and controlled processing pathways rely possibly on the action of a neuro-transmitter, dopamine; see e.g., Braver-Cohen (2000) for a model of one such process, the 'dopamine gating system.' These processes do not require relying on an 'homunculus;' see Monsell-Driver (2000).

[15]The process of activating and maintaining relevant representations in the prefrontal cortex is analogous to the process involved in working memory tasks; see Miyake-Shah (1999). Brain imaging evidence has been proposed which supports the direct role of working memory and attention in the executive function's modulation of the interplay of automatic and controlled processes in cognitive control tasks; see e.g., Engle (2000). Also, see Engel-Kane-Tuholski (1999), Just-Carpenter (1992) on the limits of the activation capacity of the cortex.

[16]The view that decision making arises from the interaction of automatic and cognitive processes, or visceral and rational states, is at least as old as the Bible. It has been exploited most notably in recent times in psychoanalytic theory where it takes the form of the *Ego* and the *Id* (see Freud (1927). A formal model was introduced in economics by Thaler-Shefrin (1981). The related work of Loewenstein

As an illustration of the behavior and of the brain processes associated to cognitive control, consider a specific cognitive control task, the Stroop task, after the experiments by Stroop in the 30's. The task consists in naming the ink color of either a conflicting word or a non-conflicting word (e.g., respectively, saying 'red' to the word 'green' written in red ink; and saying 'red' to the word 'red' written in red ink). The standard pattern which is observed in this experiment is a higher reaction time for conflicting than non-conflicting words. Moreover the reaction time is higher, in either case, than the reaction time of a simple reading task; and the reaction time of a reading task is unaffected by the ink color. Cohen-Dunbar-McClelland (1990) have developed a 'connectivist' (loosely, biologically founded)[17] cognitive control model of the Stroop task which generates the same pattern of reaction times that are observed in the experiments; see also Braver-Cohen-Servan Schreiber (1995) and Braver-Cohen (2000). In their model, word-reading is a strong association encoded in the posterior cortex, which produces a rapid automatic response. The controlled processing aspect of the task is identified in naming the ink color: color-naming is a weaker association, but it can override the stronger word-reading process if it is supported by the activation of the pre-frontal cortex to maintain the appropriate task-relevant goal by inhibiting the automatic reading association. Importantly, brain imaging data of subjects during Stroop show the sustained neural activity in the pre-frontal cortex that is consistent with this interpretation; see Miller-Cohen (2001).[18,19]

The basic postulate of this paper is that internal commitment mechanisms and self-control operate as cognitive control mechanisms in dynamic choice. We make the connection between cognitive control, internal commitment, and self-control more precise by illustrating a possible cognitive control mechanism which might induce self-control in a simple delayed gratification choice task. In the next section we will extend our model of delayed gratification choice into an analysis of a dynamic consumption-saving problem.

Consider an agent planning his optimal consumption allocation between two periods in the future. In particular, an agent at time $\tau = 0$ must choose how to distribute a

---

(1996) and Bernheim-Rangel (2001), like ours, is instead motivated by neurobiological evidence. The identification and the modelling of the neural processes responsible for cognitive control, and especially of the mechanism which modulates the differential activation of such processes, is the recent contribution of cognitive sciences which we are introducing to the study of dynamic decision making and which characterizes our approach. The foundations of our model of internal commitment and self-control lie in the explicit modelling of cognitive control processes rather than in visceral/rational dichotomy per se.

[17]See McClelland-Rumelhart (1986); also, O'Reilly (1999) for a list of principles of 'connectivist' modelling.

[18]Furthermore, patients with frontal impairment have difficulties with the Stroop task; see Cohen-Servan Schreiber (1992) and Vendrell-Junque-Pujol-Jurado-Molet-Grafman (1995).

[19]Another extensively studied task which requires cognitive control is the anti-saccade. In these experiments the interaction between automatic and controlled determinants of behavior is elicited through a task which requires the experimental subject to inhibit a powerful drive to automatically saccade to an abrupt visual cue; see e.g., Curtis-D'Esposito (2003).

given income endowment w for consumption in the future at time $t > 0$ and time $t + 1$. An agent with preferences represented by utility function $U(c)$ for consuming c units of the consumption good, and with exponential discounting at rate $\beta < 1$, would solve the following maximization problem:

$$\max_{c_t, c_{t+1}} \beta^t \left[U(c_t) + \beta U(c_{t+1}) \right] \tag{1}$$

$$s.t. \quad c_t + c_{t+1} \leq w \tag{2}$$

Let the solution to this problem be denoted by $(c^*, w - c^*)$; it represents the agent's goal or plan. When the same agent faces the same problem in the present, that is when the first component of the choice can be consumed immediately, $\tau = t$, the agent faces a different 'temporary' preference representation induced by a strong automatic association which favors immediate consumption over delayed consumption. For instance, the agent would rather consume in this case $c^I > c^*$ at time t. In so doing the agent would 'reverse' his time preferences as the delayed gratification choice becomes nearer to the present, as τ goes to t.[20] The agent's ability to delay gratification possibly results then only from internal commitment mechanisms. We postulate that such mechanisms operate as cognitive control. When the agent is given the delayed gratification choice at time $\tau = t$, an automatic process is activated which would induce him to choose c^I at time t, leaving $w - c^I$ for time $t + 1$. At time t controlled processing is also activated. It operates through actively maintaining in the frontal cortex the representation of his planned consumption choice c^*, as a goal, and possibly overriding the choice induced by automatic processing by inhibiting its activation. Inhibitory connections are activated depending on expected future rewards, $U(c^*) - U(c^I) + \beta \left[U(w - c^*) - U(w - c^I) \right]$. Since maintaining an active representation is costly, in terms of the limited activation capacity of the supervisory attention system, we postulate that inhibitory connections override the automatic processing pathway if

$$U(c^*) - U(c^I) + \beta \left[U(w - c^*) - U(w - c^I) \right] > b \tag{3}$$

for some parameter b measuring attention costs, or the costs of maintaining a representation in active memory.[21] The interpretation of b as attention costs is consistent with the

---

[20]In fact, psychologists have documented this phenomenon, called *reversal of preferences*, in several specific experimental implementations of the delayed gratification choice task; see e.g., Kirby (1997) and Kirby-Herrnstein (1995), and Ainsle (1992), (2001), Ainsle-Haslam (1992), Frederick-Loewenstein-O'Donoghue (2002), Herrnstein (1997) for comprehensive surveys. See Ainsle (1992) for an insightful discussion of the dependence of the incidence of reversal of preferences on the experimental design.

[21]This formulation is related to k-winners-take-all models of inhibitory functions in Majani-Erlason-Abu Mostafa (1989), which have been adopted e.g., by O'Reilly-Munakata (2000) to study cognitive control.

classical view in psychology that considers self-control a form of attention control.[22,23] $U(c^*) - U(c^I) + \beta \left[U(w - c^*) - U(w - c^I) \right]$ can be interpreted as a measure of the *regret* (in utility scale) the agent faces once his 'temporary' preference representation vanishes if he has chose to consume c^I.[24]
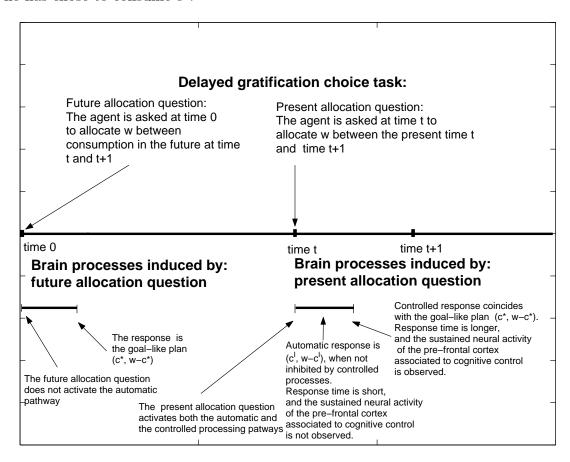

Figure 1: **Delayed Gratification: Timeline**

The neurobiological foundation of the basic postulate of this analysis, that self-control in delayed gratification choice tasks is a specific form of cognitive control has never been

---

[22]For instance, William James, concluding the analysis of 'will' in *The Principles of Psychology*, Holt, 1890, states: 'effort of attention is thus the essential phenomenon of will', and 'the difficulty [of self-control] is mental: it is that of getting the idea of the wise action to stay before your mind at all' (pg. 1167; cited in Shefrin-Thaler (1992)).

[23]Attention costs b are conceptually distinct from computational costs. In this paper we abstract from computational costs, even though they affect dynamic choice.

[24]More abstractly, the dynamic choice procedure induced by regret and attention costs in (3) can possibly be justified also axiomatically, along the lines of Gul-Pesendorfer (2001). We leave this for future work.

tested with imaging data.[25] This would require developing a 'connectivist' model of delayed gratification choice, along the lines of Cohen-Dunbar-McClelland (1990)'s model of Stroop. The delayed gratification choice task could then be implemented experimentally to induce the subjects to exercise internal commitment mechanisms that override the impulse to reverse preferences. Reaction time and imaging data from this experiment, when matched with data generated by the delayed gratification choice model, could be used to test whether cognitive control drives the operation of internal commitment mechanisms and self-control; see Figure 1 for a more detailed representation of the delayed gratification choice task experiment.

Some indirect evidence in favor of our analysis of the delayed gratification task has been collected by cognitive psychologists. Our analysis in fact, based on the limitation of the activation capacity of the supervisory attention system, predicts that self-control is harder to exercise when an agent is performing unrelated cognitive tasks simultaneously. It is therefore consistent with Shiv-Fedorikhin (1999)'s and Vohs-Heatherton (2000)'s experimental data documenting a reduction of self-control in subjects asked to perform parallel working memory tasks. Experimental treatments of delayed gratification choice tasks under differential capacity utilization of working memory would generate additional behavioral and imaging data with the power of testing our model of internal commitment and self-control.

3   Consumption-saving decisions

In this section we extend the analysis of cognitive control and delayed gratification of the previous section to study the consumption and saving behavior induced by an agent's internal commitment ability. We develop a cognitive control model to identify self-control strategies for consumption-saving behavior. As noted in the Introduction, standard models in behavioral economics ignore the internal commitment ability of the agents.

3.1   The economy

Consider a dynamic economy, with time indexed by $t = 0, 1, ..., \infty$. Let the consumer's utility for c_t units of the good at time t be denoted $U(c_t)$. The agent faces a linear production technology, and the wealth accumulation equation is

$$k_{t+1} = a_t k_t - c_t \tag{4}$$

---

[25]Applications of brain imaging methods to choice tasks in general include Rustichini-Dickhaut-Nagode-McCabe-Pardo-Smith (2003) on risk and uncertainty, McCabe-Houser-Ryan-Smith-Trouard (2001) on the theory of the mind, and Sanfey-Rilling-Aronson-Nystrom-Cohen (2003) on the ultimatum game.

where k_t and c_t denote respectively the agent's wealth and consumption at time t; and a_t is the productivity parameter at t. The productivity a_t is in general stochastic.

Assumption 1 *The productivity a_t is i.i.d., takes values in $(0, \infty])$, and has well-defined mean, $E(a) > 0$.*

At any time t the agent observes a "temptation", z_t. The effect of the temptation is to generate a 'distorted' *temporary* representation of preferences at time t of the form:

$$U\left(z_t c\right)$$

Assumption 2 *The temptation z_t is i.i.d., takes values in $[1, \infty)$, and has well-defined mean, $E(z) > 1$.*

To interpret preferences $U\left(z_t c\right)$ as subject to temptation, we assume that under this representation the perceived marginal utility of consumption at time t is higher than it is under preferences $U(c_t)$, for any c_t.

Assumption 3 *The consumer's utility for consumption, $U\left(c\right)$, is Constant Elasticity of Substitution (CES):*

$$U\left(c\right) = \frac{c^{1-\sigma}}{1-\sigma}$$

with $\sigma < 1$.

Note that, with our formulation of preferences, it is $\sigma < 1$ that guarantees that the marginal utility of consumption increases with temptations $z_t \geq 1$.[26],[27] Since the production technology is linear and preferences are CES, we restrict attention to linear consumption plans of the form

$$c_t = \lambda_t a_t k_t$$

where λ_t, the propensity to consume at time t, is the consumer's choice variable.[28] The implied accumulation equation for capital becomes

$$k_{t+1} = \left(1 - \lambda_t\right) a_t k_t$$

---

[26]If $\sigma > 1$, the utility function and the value function are negative, so a temptation that increases temporary utility would require values of $z_t \leq 1$, and an increase in future temptations should be characterized by a stochastic decrease in the distribution of $z_{t+\tau}$. While our basic analysis remains unaffected, some of our comparative static results will change if $\sigma > 1$.

[27]We model temptation as a shock to the utility function rather than as a shock to the discount rate. With CES preferences and a single commodity, as in our case, this hardly makes a difference, but the distinction is important in more general models in which temptations can affect different goods differently, e.g., in models with addictive and normal goods.

[28]This is in fact without loss of generality, as our subsequent analysis demonstrates.

3.2 Cognitive control and consumption-saving

Agents have the ability to invoke either automatic processes that are susceptible to the temptation of 'over-consuming,' or alternative control processes which are immune to such temptations, along the lines of the models of cognitive control and delayed gratification introduced in the previous section. We do not endow the agents with any external commitment mechanism, so that their consumption-saving behavior is governed exclusively by internal commitment and self-control strategies.

An agent facing a self-control problem observes z_t, the temptation he is facing at t, which determines the marginal utility of present consumption under his 'temporary' representation of preferences. Decision making arises from the interaction of automatic and controlled processing. Automatic processing produces a consumption-saving rule, given a_t and z_t, represented by a propensity to consume λ_t^I. For most of our analysis it is not important that λ_t^I solves a well defined maximization problem. We therefore only require λ_t^I to be represented by a continuous map $\lambda^I(z_t)$, increasing with the temptation z_t.[29] Following the realization of a_t and z_t, controlled processing is also initialized. It disregards the 'temporary' preference representation induced by z_t and it also produces a consumption-saving rule in the form of a propensity to consume λ_t. This consumption saving rule optimally trades off immediate consumption for future consumption but recognizes the interaction that will determine which processing pathway is active at each future time $t+\tau$, given $a_{t+\tau}$ and $z_{t+\tau}$. In particular, we assume that controlled processing operates by correctly anticipating the stochastic properties of temptations and the results of its interaction with automatic processes for consumption-saving in the future.[30]

We proceed by formally deriving the consumption-saving rule resulting from the activation of controlled processing. The controlled processing pathway first computes the future value of the consumption-saving plan, $D(a_{t+1}, k_{t+1}, z_{t+1})$ which depends on the active process at each future time $t + \tau$, given $a_{t+\tau}$ and $z_{t+\tau}$. Temptations will not be inhibited at all future times as it is costly, in terms of activation capacity, to choose a propensity to consume smaller than the one induced by automatic processing responding to temptation. At some future times t, λ_t^I may be such that some $\lambda_t \geq \lambda^I(z_t)$ will in fact be chosen.[31] As in the cognitive control and delayed gratification model in the previous section, we assume that the results of the interaction between processing pathways are determined by a 'supervisory attention system' governed by expected rewards. Suppose in particular that the automatic process is only active if the utility loss (or expected future regret) associated with the temptation is smaller than an exogenous activation

---

[29]In Appendix A we consider two simple specific algorithms for the automatic processing as an illustration.

[30]Correct anticipations could be based on reinforcement learning procedures. But see also Loewenstein-O'Donaghue-Rabin (2002) for evidence from survey data regarding a 'cold-to-hot empathy gap,' that is a projection bias in predicting future utility.

[31]Little of substance will be lost if we assume that, when the automatic pathway is not inhibited, $\lambda_t = \lambda^I(z_t)$.

cost $b(a, k)$, with the following simple functional form: $b(a, k) = b(a_t k_t)^{1-\sigma}$. (We adopt this functional form to guarantee the stationarity of the consumption-saving decision in order to simplify the problem.) In this case, $D(a_t, k_t, z_t)$ is given by:

$$
\begin{aligned}
D(a_t, k_t, z_t) &= \\
\max &\left[\begin{array}{l} \max_{\lambda \geq \lambda_t^I} U(\lambda a_t k_t) + \beta E\left[D\left(a_{t+1}, (1-\lambda) a_t k_t, z_{t+1}\right)\right], \\ \max_\lambda U\left(\lambda a_t k_t\right) + \beta E\left[D\left(a_{t+1}, (1-\lambda) a_t k_t, z_{t+1}\right)\right] - b(a_t k_t)^{1-\sigma} \end{array} \right] \\
&\text{with } \lambda_t^I = \lambda^I(z_t)
\end{aligned} \tag{5}
$$

Given the future value of the consumption plan, $D\left(a_{t+1}, k_{t+1}, z_{t+1}\right)$, the controlled processing pathway computes the desired consumption-saving rule as the propensity to consume λ_t which solves:

$$
\max_\lambda U\left(\lambda a_t k_t\right) + \beta E\left[D\left(a_{t+1}, (1-\lambda) a_t k_t, z_{t+1}\right)\right] \tag{6}
$$

The resulting propensity to consume is independent of z_t; let it be denoted $\lambda^E(a_t, k_t)$.

As we noted earlier, expected rewards determine the results of the interaction between the automatic and the controlled processes. This interaction, implicit in the determination of $D(a_t, k_t, z_t)$ in (3-5), can be represented simply as follows. Given $\lambda_t^I = \lambda^I(z_t)$ and $D(a_t, k_t, z_t)$, the utility loss (expected future regret) associated with the temptation z_t at time t, is

$$
\begin{aligned}
R(a_t, k_t, z_t) &= \left[\max_\lambda U\left(\lambda a_t k_t\right) + \beta E\left[D\left(a_{t+1}, (1-\lambda) a_t k_t, z_{t+1}\right)\right]\right] \\
&\quad - \left[\max_{\lambda \geq \lambda_t^I} U(\lambda a_t k_t) + \beta E\left[D\left(a_{t+1}, (1-\lambda) a_t k_t, z_{t+1}\right)\right]\right]
\end{aligned}
$$

Inhibitory controls activate controlled processing if

$$
R(a_t, k_t, z_t) > b(a_t k_t)^{1-\sigma}
$$

In summary, the present bias in the model derives from the stochastic temptation that affects the computations of automatic processing. Self-control at time t coincides with disregarding temptation z_t in the decision process. It requires the active maintenance of a goal-like representation of a consumption-saving rule which is independent of the temptation z_t. Such a representation is maintained by the force of the inhibitory connections linking the reward predictions and active representation. The 'supervisory attention system' modulates the updating of the active representations by the activation of inhibitory connections which are stronger the higher is the prediction of regret given (a_t, k_t, z_t).

3.2.1 Characterization

In this section we characterize the consumption-saving behavior of an agent in our cognitive control model. Given $\lambda^I(z_t)$ we solve for the future value of the consumption-saving plan, $D(a_t, k_t, z_t)$, and the consumption-saving plan associated with controlled processing, $\lambda^E(a_t, k_t)$. The agent's behavior is then determined at each time t by the interaction between processing pathways: the agent's propensity to consume is $\max\{\lambda^E(a_t, k_t), \lambda^I(z_t)\}$ when he expects a limited future utility loss (regret), $R(a_t, k_t, z_t) \leq b(a_t k_t)^{1-\sigma}$, while the temptation z_t is inhibited and the propensity to consume is $\lambda^E(a_t, , k_t)$ if $R(a_t, k_t, z_t) > b(a_t k_t)^{1-\sigma}$.

Given $\lambda^I(z_t)$, each agent's consumption-saving plan is characterized by the policy function of the dynamic programming problem (5).

Proposition 1 *The value function $D(a_t, k_t, z_t)$ defined by problem (5) exists. The consumption-saving rule associated with controlled processing, $\lambda^E(a_t, k_t)$, is in fact a constant, λ^E. Moreover, there exist a unique policy function of problem (5), $\lambda(a_t, k_t, z_t)$, which has the following properties:* i) *it is independent of (a_t, k_t), that is $\lambda(a_t, k_t, z_t) = \lambda(z_t)$;* ii) *it has a cut-off property, that is, there exists a $\overline{\lambda}$ such that*

$$\lambda(z_t) = \begin{cases} \max\{\lambda^E, \lambda^I(z_t)\} & for \quad \lambda^I(z_t) \leq \overline{\lambda} \\ \lambda^E & else \end{cases} \tag{7}$$

An alternative representation of the policy function of problem (5) can be derived in which automatic processing is inhibited at a time t for large enough realized temptations z_t. This is an immediate corollary of Proposition 1 and of our assumption that $\lambda^I(z_t)$ increases with z_t, that is, that the propensity to consume associated with a automatic processing increases with the intensity of the realized temptation.

Proposition 2 *There exist a $\overline{z}$ such that*

$$\lambda(z_t) = \begin{cases} \max\{\lambda^E, \lambda^I(z_t)\} & for \quad z_t \leq \overline{z} \\ \lambda^E & else \end{cases} \tag{8}$$

The behavior of an agent facing conflicting preference representations over his consumption-saving choice in our cognitive model can be quite simply summarized: He actively maintains a simple consumption-saving goal, a propensity to consume out of wealth which is independent of any realized temptation, and is equal to λ^E. At times the agent allows temptations to affect his consumption-saving behavior by letting the impulsive choice induced by automatic processing $\lambda^I(z_t)$ prevail, if this choice does not perturb his underlying consumption-saving plan too much and therefore does not have large permanent

effects on his prescribed wealth accumulation. In particular, controlled processing inhibits automatic processing when temptations are large enough.[32]

It is important to notice that our specific characterizations depend in a crucial manner on our assumptions regarding attention costs $b(a_t k_t)^{1-\sigma}$. As we noted, the specific functional form depending on a_t and k_t is adopted to simplify the computations, by maintaining homogeneity with the CES preferences. The implicit assumption that b is constant, and in particular independent of the realized temptation, z_t, is however substantial. While this is a natural assumption if such costs are interpreted literally as attention costs, in principle it is important to explore different formulations that relate costs to the size of the temptations. In particular, if costs are small for $z = 0$ and increasing in z, some small temptations may also be inhibited.

3.2.2 Properties of cognitive control

Consider different environments in terms of the stochastic process of temptations. In particular, we identify more tempting environments with a first order stochastic dominance increase in the distribution of future temptations z_τ, for $\tau > t$;[33] that is, essentially a shift of some mass from lower realization of z_τ's into higher realization of z_τ's.[34]

Proposition 3 *Let the random variables a_τ and z_τ be independent, for all $\tau > t$. i) The propensity to consume associated with controlled processing, λ^E, increases with an increase in the first order dominance sense of the distribution of future temptations z_τ, $\tau > t$. ii) The cut-off $\overline{\lambda}$ is decreasing with an infinitesimal increase in the first order dominance sense of the distribution of z_τ, $\tau > t$.*

---

[32]When temptations are small however the agent will choose λ^E without the need to inhibit automatic processing (and hence to incur the related attention costs given by b) only if $\lambda^I(z_t) < \lambda^E$. This may occur only for specific forms of automatic processing that can result in too much saving; otherwise, and more naturally, $\lambda^I(z_t) > \lambda^E$ for any z_t; see the automatic processing examples in Appendix A.

[33]Let f and f' denote two probability densities on a compact subset of $\Re$, X, and let F and F' denote the associated cumulative functions. The density f' dominates in the first order stochastic sense the density f if

$$F'(x) \leq F(x), \ \forall x \in X.$$

Moreover, fix a density f' which dominates f in the first order stochastic sense, and consider the distribution obtained by mixing $f(x)$ with $f'(x)$:

$$g(x) = (1 - \alpha)f(x) + \alpha f'(x)$$

By an infinitesimal increase in the first order dominance sense in the distribution of x we mean an infinitesimal increase $d\alpha > 0$ evaluated at $\alpha = 0$.

[34]In the following propositions we keep the map $\lambda^I(z)$ fixed. The results are more general though and could be extended to automatic processing mechanisms which react to different distributions of temptation. In particular the propositions hold for both the mechanisms studied in Appendix A.

The intuition for the effects of an increase in the first order dominance sense in the distribution of z_t hinges on the fact that the expected future value of the consumption-saving program represents the marginal value of savings. If a change in the distribution of temptations has the effect of decreasing the expected future value of the consumption-saving program, then at the margin an agent, independently of whether he exercises self-control or not, will save less and consume more in the present. This is in fact the effect of an increase in the first order sense of the distribution of z_t: the value of the program is weakly decreasing in z_t and hence an increase in the distribution of z_t in the first order sense, shifts probability mass from realizations of temptations associated with higher values of the program to realizations associated with lower values of the program, thereby decreasing its expected value.[35] But in our model an agent counterbalances the lower savings rate associated with controlled processing with a more stringent rule regarding the conditions under which temptations are not suppressed and automatic choice not inhibited. After an increase in the first order dominance sense of the distribution of z_t, the cost of inhibiting automatic processing is unchanged and equal to b, while the value of inhibition is on average higher, since the distribution of z_t has shifted towards higher realizations of z_t.[36]

Drawing on the implications of Proposition 3, we note that an increase in the first order dominance sense of the distribution of z_t increases by definition the mass of the distribution of z_t on $z > \overline{z}$, for any $\overline{z}$. Furthermore, an increase in the first order dominance sense in the distribution of z_t, generated by a shift of mass from z such that $\lambda^I(z) > \lambda^E$, decreases the cut-off $\overline{\lambda}$. As a consequence, $\overline{z}$ decreases. We conclude then that an increase in the first order dominance sense of the distribution of z_t increases the probability that self-control is exercised and automatic choice is inhibited. On the other hand, a local (infinitesimal) increase in the first order dominance sense of the distribution of z_t increases λ^E, i.e., the consumption when self-control is exercised and automatic processing inhibited. We conclude that an agent facing larger temptations in the future reacts by exercising self-control more often but at the same time by consuming a higher fraction of his wealth even while controlling himself.

---

[35]Recall that we have assumed $\sigma < 1$. Results in this and the next section depend on this assumption. This is because the savings rate is increasing in the rate of return if $\sigma < 1$, while it is decreasing if $\sigma > 1$ (and constant in the log case, $\sigma = 1$. In particular a decline in the future value of the program due to future temptations, which is similar to reductions of the rate of return, will induce larger, not smaller saving rates, if $\sigma > 1$. Even if $\sigma \geq 1$, however, controlled processing will inhibit savings rates under automatic processing if they are too low relative to the preferred choice, $1 - \lambda^E$, and our analysis of the inhibition of excessive consumption binges induced by temptations will hold with minor modifications in the case $\sigma \geq 1$.

[36]In fact, a countervailing effect must be taken into account: the value of inhibiting automatic processing is reduced by the increase in λ^E due to the same increase, in the first order dominance sense, of the distribution of future temptations z_τ, $\tau > t$ (Proposition 3-i). But this effect is second order for infinitesimal changes in the distribution of z_t by the Envelope Theorem.

Our cognitive control model allows us also to study the dependence of consumption-savings behavior on differences in the internal psychological characteristics of an agent, e.g., cognitive abilities like setting goals and controlling attention, affecting consumption-saving behavior. As already noted, different 'propensities to plan' have been documented by Ameriks-Caplin-Leahy (2002) with survey data on retirement savings. Also, different cognitive abilities have been extensively documented in the psychological literature; see e.g., Baumeister-Heatherton-Tice (1994) for a survey. In particular, in our set-up we can study the comparative statics of consumption-saving behavior with respect to the attention cost parameter b which determines an agent's cognitive ability to inhibit automatic impulsive preference representation, and hence to self-control: an increase in b increases the cost of inhibiting automatic processing at any time t, and hence the cost of exercising self-control.

Proposition 4 *Let the random variables a_τ and z_τ be independent, for all $\tau > t$.* i) *The propensity to consume associated with controlled processing, λ^E, increases with an increase in b.* ii) *The cut-off $\overline{\lambda}$ increases in b.*

Not surprisingly, an increase in attention costs b has the effect of increasing the cut-off $\overline{\lambda}$, that is, of rendering it less stringent. Moreover, an increase in b reduces the expected future value of the consumption-saving program, by making it more costly to exercise self-control, and hence it reduces the marginal value of saving; consequently, a higher b induces a larger propensity to consume associated to controlled processing.

Finally, we address the important issue of the effects of the complexity of savings goals on consumption-savings behavior. The behavior of an agent facing conflicting preference representations over his consumption-saving choice in our model, as we noted, involves actively maintaining a simple consumption-saving goal. Such a goal consists of a propensity to consume out of wealth which is independent from any realized temptation. Psychologists constantly remark that the *complexity* of the goals individuals set for themselves affects their ability to self-regulate and exercise self-control in particular tasks.[37] The simple formulation of the agent problem we have adopted however, with linear production technology and CES preferences, implies that the consumption-saving goal is extremely simple: it is constant over time, as it is independent of the realization of the production shock a_t. To study the issue of complexity of the goal agents set for themselves, we need to examine instead a more general formulation of the model, which potentially gives rise to more complex consumption-savings plans in the event of self-control. As a way of illustration consider the following formulation of technology, leaving preferences unchanged:

$$k_{t+1} = R_t \left(a_t k_t - c_t \right), \ R_t, a_t \geq 0 \tag{9}$$

---

[37]The books by Baumeister-Heatherton-Tice (1994), and Gollwitzer-Bargh (1996), for instance, discuss the rich literature on the topic.

In this formulation the shock R_t acts on net wealth $k_t - c_t$, and therefore takes the interpretation of a rate of return on saving at t (a_t is instead a productivity shock, as in the case of the technology studied in the previous section, equation (4); we assume it independent of R_t). The novel feature of this formulation is that he value of controlling any temptation is random, and proportional to the realization of R_t : If for instance the return on saving is small, R_t is small, self-control is of little use. As a consequence, the consumption-saving plan depends on R_t; let it be denoted $\lambda(z_t, R_t)$. Let also $\lambda^I(z_t, R_t)$ and $\lambda^E(R_t)$, denote the propensities to consume associated, respectively, with the automatic and the controlled pathways; let finally $\overline{\lambda}(R_t)$ denote the cut-off which characterizes $\lambda(z_t, R_t)$.

Therefore in this environment we can study the issue of the complexity of the goal $\lambda^E(R_t)$, with respect to any simpler goal represented by a constant consumption-saving plan over time, that is a plan independent of R_t. Suppose in fact that the activation cost parameter, b, decreases with the complexity of the goal that is to be maintained active in conscious memory. In particular, we interpret this to mean that activation costs are lower to maintain a constant consumption-saving rule, $\lambda^{E,simple}$, than they are to maintain a fully contingent plan $\lambda^E(R_t)$. Here we take the constant plan $\lambda^{E,simple}$ to coincide with the optimal consumption-saving plan associated with cognitive control under the restriction that $\lambda^{E,simple}$ is independent of R_t at any time t.[38] Our objective is to characterize conditions for the parameters under which an agent would gain from setting the simpler constant goal rather than the 'complex' goal that is contingent on the state of the technology, R_t.

Let the activation cost associated with the simple plan be denoted b^{simple}, and let the difference in the cost parameters between the simple and complex goal be denoted by Δb.

Proposition 5 *A simpler constant consumption-saving plan $\lambda^{E,simple}$ tends to be preferred to the complex plan $\lambda^E(R_t)$ if in the limit, and other things equal,* i) *b^{simple} is small and Δb large enough,* ii) *the mean of R_t is small enough, and finally if* iii) *the mean of z_t as well as Δb are large enough.*

The simple consumption-saving plan is preferred to the complex plan, not surprisingly, if it is easy to keep it active, and much easier than maintaining the complex plan. More interestingly, the simple plan is preferred if the mean of the stochastic rate of return, $E(R_t)$, is small enough, or close to 0. In this case, since the support of the rate of return shocks is $[0, \infty)$, the variance of R_t also tends to 0 and hence rate of return in the limit is degenerate, and concentrated on 0. But in this case self-control is useless, and it is

---

[38]In fact, an agent could learn and encode a simple un-contingent plan as an automatic process. See Miller-Cohen (2001) and Bownds (1999) for some evidence and discussions on plasticity of the brain and changes of the representational content of automatic and controlled processing; see also Gollwitzer (1999) for psychological experiments aiming at eliciting automatic reactions in planning.

a dominant choice for the agent to consume all of his wealth each period. Therefore, the utility gain of conditioning the consumption-saving plan on the realization of R_t vanishes. The simple plan is also preferred if the mean of the stochastic process of temptations grows large. This is because when temptations are large enough, in the limit, the complex plan will optimally induce inhibition of the automatic processing all the times, independently of R_t, and this behavior can also be induced by a simple plan. (The condition on Δb is required since the savings in terms of attention costs associated with the simple plan must of course more than compensate the loss of utility from the adoption of the non-contingent plan by itself, once inhibition is guaranteed at all times.)

3.3 Benchmarks: Exponential maximizers and intra-personal dynamic games

Our model of internal commitment and self-control nests two important alternative models of consumption-saving, the Life Cycle/Permanent Income model with exponential discounting and the behavioral model of the strategic interaction of multiple successive selves. They correspond, respectively, to the extreme cases in which $b = 0$ and the agent can inhibit temptations at no costs, and in which $b = \infty$ and no temptation can be inhibited.

We study these alternative models in turn. Consider an agent who never faces temptations and self-control problems, that is, a Life Cycle/Permanent Income exponential discounter. In our economy such an agent will choose the constant consumption-saving plan λ^*, determined as the solution of the following recursive maximization problem:

$$V\left(a_t, k_t\right) = \max_{\lambda} \ (1 - \sigma)^{-1} \left(\lambda a_t k_t\right)^{1-\sigma} + \beta E V\left(a_{t+1}, (1 - \lambda) a_t k_t\right) \qquad (10)$$

(The closed form solution for λ^* is derived in the Appendix B; it corresponds to the special case with $z_t = 1$, $E z_{t+1} = 1$ of the result of Lemma A.1.)

It is easy to see that, in our model, $D\left(a_t, k_t, z_t\right)$ converges to $V\left(a_t, k_t\right)$ if attention costs b converge to 0, and the agents can inhibit temptations at no costs, $\lambda^E = \lambda^*$ and $\overline{\lambda} \leq \lambda^E$.

It is natural to assume that that

$$\lambda^I(z_t) \geq \lambda^*, \ \forall z_t \geq 1,$$

that is, consumption-saving plan associated with the automatic pathway to imply a propensity to consume which is in any case larger than or equal to the propensity to consume of an agent with no self-control problems. In this case, if attention costs are positive, $b > 0$, Proposition 2 can be extended to show that

$$\lambda^E > \lambda^* \qquad (11)$$

The consumption-saving goal determined by controlled processing requires more consumption and less savings than is optimal from the point of view of a Life Cycle/Permanent Income agent who never faces temptations. The intuition for this result hinges once again on the expected future value of the consumption-saving program, which at the margin represents the value of savings. The expectation of self-control problems in the future has the effect of depressing the expected future value of the consumption-saving program, and hence at the margin it induces less saving and more consumption in the present.

Consider instead the decision problem of an agent who does face self-control problems, in the sense that he perceives a strategic interaction with future selves with different preference orderings, and plays a Markov Perfect Nash equilibrium of the dynamic game. As already noted this represents the standard approach of behavioral economics, as e.g., in Laibson (1996) and O'Donoghue-Rabin (1999). We extend it in the following to account for our stochastic economic environment, by letting the agent's preferences at time t depend on the realization of the temptation z_t, but not on any future temptations. The agent however will, at time t, anticipate that preferences of his future selves will depend on the future temptations.

Formally, the agent's behavior in equilibrium is determined as a consumption-saving rule $\lambda^M(z_t)$ solving the following fixed point condition:

$$\lambda^M(z_t) = argmax_\lambda (1-\sigma)^{-1} (z_t \lambda a_t k_t)^{1-\sigma} + EV_{\lambda^M(z)}(a_{t+1}, (1-\lambda)a_t k_t, z_{t+1}) \qquad (12)$$

where $V_{\lambda(z)}(a_t, k_t, z_t)$, the value at t of present and future consumption induced by an arbitrary consumption-saving rule $\lambda(z)$, is defined by

$$V_{\lambda(z)}(a_t, k_t, z_t) = (1-\sigma)^{-1} (\lambda(z_t)a_t k_t)^{1-\sigma} + E \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} (\lambda(z_\tau)a_\tau k_\tau)^{1-\sigma} \qquad (13)$$

From the point of view of the agent's time t self, the value of present consumption is directly affected by the temptation z_t, while the the expected value of future consumption, $EV_{\lambda(z)}(a_{t+1}, k_{t+1}, z_{t+1})$, is affected by future temptations z_τ only through the expectation of future choices $\lambda(z_\tau)$.

Although we did not obtain a closed form solution for $\lambda^M(z_t)$, the following result provides a simple characterization of a Markov Perfect Nash equilibrium consumption-saving rule.

Proposition 6 *With respect to the consumption-saving plan of an exponential maximizer, at a Markov Perfect Nash equilibrium the propensity to consume out of wealth is larger: $\lambda^M(z_t) > \lambda^*$, for any z_t. Moreover, it is increasing in z_t.*

At a Markov Perfect Nash equilibrium of the game of successive selves, even though the preferences of agent t are independent of future temptations, an agent anticipates

that his future selves will in fact face stochastic temptations and will not exercise self-control: he expects from all future selves the same behavioral rule he himself adopts, and in equilibrium he sets his present consumption-saving rule accordingly. The expected future value of the consumption-saving program at the margin represents the value of savings. The expectation of self-control problems in the future has the effect of depressing the expected future value of the consumption-saving program, and hence at the margin an agent facing self-control problems will save less and consume more in the present. Even if the agent faces no temptation at t, that is, $z_t = 1$, the expectation of future temptations not controlled by his future selves reduces his incentives to save at time t and hence induces a larger propensity to consume out of wealth. In fact, at each time t the consumption-saving rule depends on the time t realization of the temptation, z_t; and since at equilibrium the agent never exercises self-control and always succumbs to the temptation, the higher the temptation the more he consumes.

It is immediate to see that, in our model, when $b = \infty$ the agents cannot ever inhibit temptations and his consumption-saving plan is determined by the content of the automatic processing pathway. It is then natural to consider the case where $\lambda^I(z_t) = \lambda^M(z_t)$, and the propensity to consume associated to the automatic pathway coincides with the Markov Perfect Nash equilibrium of the game of successive selves given by the solution of problem (12-13). This is in fact one of the cases studied in detail in Appendix A. We can now compare the behavior induced by our formulation of self-control with the behavior induced by the Markov Perfect Nash equilibrium of the game of multiple successive selves. First of all, we stress that in our model , as long as attention costs b are not infinitely high, self-control has the natural effect of limiting consumption binges driven by present and expected future temptations.

Proposition 7 *The propensity to consume induced by controlled processing, λ^E, is smaller than the propensity to consume induced by the Markov Perfect Nash equilibrium of the game of multiple successive selves, $\lambda^M(z_t)$, for any realization of the temptation z_t.*

In particular, even if no temptation is realized at time t, that is, $z_t = 1$, the savings rate implied by the Markov Perfect Nash equilibrium is lower than the savings rate implied by controlled processing. Under controlled processing agents rationally expect to exercise self-control in the future and to inhibit large temptations; as a consequence the future value of an extra unit of wealth at the margin, as of time t, is larger than at the Markov Perfect Nash equilibrium of the game of successive selves, and so the agent's incentive to save is larger as well.

3.4 Testing Against Alternative Models

Besides implying higher savings rates, the consumption-saving implications of our self-control model can be formally distinguished from those associated with the Life Cycle/Permanent Income model and those of the the Markov Perfect Nash equilibrium of

the game of multiple successive selves, even if the stochastic process driving temptations is hardly directly identified. In this section we will discuss in some detail the existing empirical evidence on consumption and savings, and argue that it provides indirect evidence in favor of our model of consumption-saving with respect to the benchmark models of Life Cycle/Permanent Income and intra-personal dynamic games.

Consider an agent who expects to be hit by an income shock in the future, e.g., a windfall gain like an unexpected wage increase, a tax rebate, or an insurance payout. If the agent is an exponential maximizer and is not liquidity constrained, as in the standard Life Cycle/Permanent Income theory of consumption, he will adjust his consumption/saving plan at the moment he learns of the shock, and no change in consumption will be observed when the agents actually receives the windfall gain. This implication of the Life Cycle/Permanent Income theory has been extensively tested with individual consumption data. The failure of this implication of the standard model is referred to as *excess sensitivity of consumption.* Consider now the identifying assumption that income shocks are correlated with temptations, in the sense that receiving a windfall gain would induce the agent to consume above his plan, unless he exercises self-control. Then, according to both our cognitive model and the intra-personal dynamic game model, we should observe some excess sensitivity of consumption. Moreover, according to our cognitive model, we should observe a propensity to consume off a windfall gain at the moment it is received which is higher when the gain is small than when it is large. In fact, we should observe no excess sensitivity for large enough shocks.

A large evidence documents excess sensitivity of consumption out of windfall gains, even after controlling for liquidity constraints;[39] see Browning-Lusardi (1996) for an excellent survey. More specifically, excess sensitivity is in fact large when windfall gains are small. Average propensities to consume of the order of 60 to 90 percent have been estimated, for instance, by Parker (1999) for changes in Social Security taxes withholdings, by Souleles (1999) for yearly IRS tax refunds, by Souleles (2002) for the Reagan tax cuts of the early $80's$,[40] and by Wilcox (1989) for Social Security benefits.[41] Much smaller propensities to consume off of windfall gains are estimated though when gains

---

[39]Results are instead mixed when expected income shocks are identified as orthogonal components of income processes. Consistently with our interpretation of excess sensitivity, in this case gains are arguably less clearly associated with temptations. Also, excess sensitivity does not appear when expected income shocks are negative; see for instance Souleles (2000) on tuition expenditures.

[40]But see Slemrod-Shapiro (2002) for much smaller estimates of the consumption effects of Bush's tax cut of 2001.

[41]Interestingly, there is some evidence that agents change their consumption plans when the gain is realized, as our model predicts: according to a New York Times/CBS News poll in May '82 agents the average propensity to consume the second phase of the Reagan tax cuts in the agents' plan was about 50 percent, while the actual propensity to consume turned out well above 80 percent in Souleles (2002) estimates. Also, the pattern of consumption after windfall gains is well in accord with a model of temptation, as expenditures are concentrated in goods like entertainment, personal care, apparel, services, but not e.g., food; see Parker (1999).

are larger: Kremin (1961) and Landsberger (1966) study Germany's restitution payments to Israeli after World War II and document propensities to consume close to 200 percent for small payments (about 1 monthly income) and as small as 20 percent for large payments (several years of income). Finally, when the payments are large, to the point of representing the main component of permanent income as in the case of unemployment insurance benefits, excess sensitivity disappears for agents who are not liquidity constrained (Browning-Crossley (2001)). Consistently, Choi-Laibson-Madrian-Metrick (2003) study the comovement of savings and unexpected wealth shocks in large sample of $401(k)$ accounts. They show that the propensity to consume out of wealth is decreasing in the size of unexpected wealth shocks.

Another important class of empirical implications that distinguishes our cognitive model from the Life Cycle/Permanent Income and dynamic game benchmarks regards portfolio allocations, and asset prices. Both our model and the dynamic game model predict that agents will allocate part of their wealth into illiquid asset, as a form of external commitment against temptations of over-consumption.[42] In particular, in the intra-personal dynamic game model this is the only form of self-control that the agents can adopt. As a consequence, the model predicts that illiquid assets should pay a negative premium (a lower return) than liquid assets. In our model illiquid assets allow agents to save on the cost of their (psychologically costly and imperfect) internal commitment strategies.Our model predicts therefore that agents would invest in such assets only when they yield a positive or a small negative premium, and hence that we should not observe a high negative premium in equilibrium.

Consistently with our model, it appears that illiquid securities pay a positive and quite sizeable return premium in asset market data; see e.g., Amihud-Mendelson (1986), Brennan-Chordia-Subrahmanyam (1998), and Pastor-Stambaugh (2001). Pastor-Stambaugh (2001), for instance, estimates a 7.5% return premium for stocks with high sensitivity to liquidity. Also[43], estimates of the return premium on educational investments, arguably the most illiquid assets, range from -2 to 7 percent. Finally, individuals' private contributions to retirement accounts also show a pattern consistent with our model. Individual Retirement Arrangement (IRA) accounts constitute a perfect external commitment asset.[44] While contributions to IRA accounts have grown rapidly in the period 1982-85, they have immediately declined after the 1986 tax reform that has limited their tax deductibility; see Venti-Wise (1987a) and Poterba-Venti-Wise (2001), especially Figure 5a.[45]

---

[42]Angeletos-Laibson-Repetto-Tobacman (2001) argue that the adoption of external commitment strategies to control consumption are important to explain the empirically observed household holdings of large illiquid assets simultaneously with costly liabilities in the U.S.

[43]This argument is directly borrowed from Kocherlakota (2001).

[44]IRA accounts have been introduced in 1982 as part of a government plan to encourage savings. Agents investing in IRA accounts (up to a fixed amount) face favorable tax treatment but are penalized for early withdrawals (before the age of 59 $\frac{1}{2}$), and for borrowing against the content of the accounts.

[45]Consistently with our model, agents seem to revert to illiquid assets with low return especially in

Finally, we consider the important implication of our model that agents will tend to adopt simple consumption-saving rules, prescribing a saving goal which is not too sensitive to negative income or productivity shocks. In fact, the evidence shows that agents only rarely reverse their saving plans, e.g., by borrowing from their home equity, or from their life insurance accounts: Venti-Wise (1987b) and Manchester-Poterba (1988) document that second mortgages are almost exclusively taken for home improvement investments, and Warshawski (1987) shows that only about 10 percent of life insurance accounts have been drawn upon.

4   Conclusions

We interpret our theoretical study of dynamic choice as introducing the functions of cognitive control in behavioral economics, by associating cognitive control with internal psychological commitment mechanisms and self-control. By considering only Markovian strategies of a game between successive selves, the behavioral economics literature implicitly models agents as lacking any form of internal psychological commitment or self-control in consumption. But only when their frontal cortex is lesioned do agents display no cognitive control. Patients with lesions in the frontal lobes display odd and impulsive behavior, they are unable to adapt to social life and conventions, and therefore hardly represent the natural object of economic analysis.[46]

While the relationship we draw from cognitive control to internal commitment and self-control is speculative at this point, we indicate how it can be tested with experimental and brain imaging data. When we apply our cognitive model of self-control to the study of dynamic consumption-saving behavior we find that it is characterized by a simple consumption-saving goal and a simple rule for invoking control processes to inhibit impulses of over-consumption and implement the consumption-saving goal. Such a rule implies that only relatively small deviations from the consumption-saving plan are allowed. While a systematic study of individual consumption-saving data is outside the scope of the present paper, our analysis of the available empirical literature on excess sensitivity of consumption clearly supports these implications of our model.

---

the context of small frequent temptations, as in the case of Christmas clubs; see Elster (1979).

[46]See Bechara-Damasio-Damasio-Anderson (1994) and Bechara-Tranel-Damasio-Damasio (1996) for the clinical analysis of behavior of frontally damaged patients.

Appendix A: Automatic Processing

4.0.1 Automatic processing

We consider by way of example two different possible mechanisms for automatic processing which satisfy the requirements imposed in Section 3.2 on $\lambda^I(z_t)$.

The two automatic processing mechanisms we study are characterized by different degrees of sophistication, in terms of their implicit anticipation of behavior. The first specification of the automatic process is related to the myopic solution to the game of successive selves introduced by O'Donoghue-Rabin (1999), once modified to account for the stochastic environment we study. The second specification of the automatic process is associated with the Markov Perfect Nash equilibrium of the game of successive selves studied by Laibson (1996), O'Donoghue-Rabin (1999) and many others.

Consider first the unsophisticated (myopic) specification for automatic process. In this case we postulate $\lambda^I(z_t)$ to solve the following recursive maximization problem:

$$V\left(a_t, k_t, z_t\right) = \max_\lambda \ (1-\sigma)^{-1} \left(z_t \lambda a_t k_t\right)^{1-\sigma} + \beta EV\left(a_{t+1}, (1-\lambda)a_t k_t\right) \qquad (14)$$

where $V\left(a_t, k_t\right)$ is defined by (10), and corresponds to the value of the consumption-saving problem of an agent not facing any self-control problem. In this formulation the automatic processing pathway is hit by a temptation z_t at t and computes the consumption-saving plan under the implicit (incorrect) assumption that no temptation will hit the agent in the future. It is easily checked that the solution satisfies all the requirements on $\lambda^I(z_t)$ imposed in Section 3.2: in particular, it is increasing in z_t; see Lemma A.1 in the Appendix for the closed form solution.

The second specification of automatic processing that we consider, which we study in the text, sets $\lambda^I(z_t) = \lambda^M(z_t)$, the Markov Perfect Nash equilibrium of the game of successive selves given by the solution of problem (12-13). In this case, automatic processing is more sophisticated, and anticipates the equilibrium choices of future automatic processing. It still is not sophisticated in another dimension, in that it does not anticipate the inhibitory activity of controlled processing, and hence it does not foresee any self-control ability for the decision-making agent.

We can now compare the consumption-saving plans represented by these two example automatic processing mechanisms.

Proposition 8 *The propensity to consume associated to automatic processing $\lambda^I(z_t)$ is smaller when determined by the myopic mechanism (14) than when determined as the Markov Perfect Nash equilibrium of the game of successive selves, (12-13).*

Moreover, i) *if $\lambda^I(z_t)$ is determined by (14), $\lambda^I(z_t) < \lambda^E$ for small enough realizations of z_t; while* ii) *if $\lambda^I(z_t)$ is determined by (12-13), $\lambda^I(z_t) > \lambda^E$ for all realizations of z_t.*

The myopic automatic processing mechanism in (14), by not anticipating future temptations, and hence by valuing the future relatively more than the more sophisticated mechanism in (12-13), is myopically induced to save more for the future. Moreover, $i)$ indicates that, if the current temptation z_t is small enough, myopic automatic processing might even be induced to save more than controlled processing. (In particular, this is true in the extreme case when the agent is not hit by a temptation at t. In this case the myopic automatic process would induce the same saving rate of an exponential maximizer, $\lambda^I(1) = \lambda^*$.) In this instance the agent will choose λ^E.

The sophisticated automatic processing mechanism in (12-13) instead, by anticipating future temptations and the associated lack of self-control of his future selves, values the future relatively little and is therefore induced to save less than myopic automatic processing and, as $ii)$ indicates, less that controlled processing, independently of the current realization of the temptation, z_t.[47]

Finally, we can compare the agent's propensities to consume induced by controlled processing when temptations are inhibited, λ^E, for either of the two different automatic processing mechanisms.[48]

Proposition 9 *The propensity to consume associated with controlled processing, λ^E, is lower when automatic processing is myopic and is determined by (14), than it is when automatic processing is determined by the solution of problem (12-13), that is, when it is governed by the Markov Perfect Nash equilibrium of the game of successive selves.*

The intuition for this result is straightforward. As indicated in Proposition 8, for any realization z_t, $\lambda^M(z_t)$ is higher than the propensity to consume implied by the (myopic) solution of (14). From the point of view of controlled processing, therefore, under the Markov Perfect Nash automatic processing, the value of the consumption-saving problem in the future is lower, and hence the propensity to consume associated with controlled processing is higher.

Appendix B: Proofs

In this Appendix we consider for simplicity an economy with a deterministic technology, $a_t = a > 0$, for any t. All proofs generalize to the stochastic case under Assumption 1.

---

[47]As a consequence, the cut-off rule governing the consumption-saving behavior of the agent is such that

$$\lambda(z_t) = \begin{cases} \lambda^I(z_t) & \text{for} \quad \lambda^I(z_t) \leq \overline{\lambda} \\ \lambda^E & \text{else} \end{cases}$$

[48]Comparison of the cut-offs for the two mechanisms leads to ambiguous results.

We first prove two lemmata. The first gives a closed form solution of the general consumption-saving maximization problem with stochastic temptations. It is referred to in the text. The second Lemma is used as a crucial component in the proofs of the propositions.

Let λ_t denote the solution of the following recursive problem:

$$V(k_t, z_t) = \max_{\lambda} \ (1-\sigma)^{-1} (z_t \lambda a k_t)^{1-\sigma} + \beta E V((1-\lambda)ak_t, z_{t+1}) \tag{15}$$

Let $\tilde{z}_t = z_t^{\frac{\sigma-1}{\sigma}}$, and $\gamma \equiv \beta^{\frac{-1}{\sigma}} \left(a^{-\frac{\sigma-1}{\sigma}}\right)$.

Lemma A.1: The solution of the maximization problem (15), λ_t, is:

$$\lambda_t = \frac{1}{1 + \gamma^{-1}\tilde{z}_t E(\tilde{z}_{t+1})^{-1} + \gamma^{-1}\tilde{z}_t E(\tilde{z}_{t+1})^{-1} E\sum_{s=t+1}^{s} \Pi_{t+1}^{s}\gamma^{t-s}} \tag{16}$$

Proof. The first order conditions of the maximization problems are:

$$z_t (z_t c_t)^{-\sigma} = \beta E V_1 (ak_t - c_t, z_{t+1})$$

$$V_1(k_t, z_t) = a\beta E V_1(ak_t - c_t, z_{t+1}) = a(z_t c_t)^{-\sigma} z_t,$$

and hence

$$z_t (z_t c_t)^{-\sigma} = a\beta E (z_{t+1}c_{t+1})^{-\sigma} z_{t+1} \tag{17}$$

Let $c_t = \lambda_t a k_t$. We can then write (17) as

$$z_t^{\frac{-1}{\sigma}} (z_t \lambda_t) = (a\beta)^{\frac{-1}{\sigma}} (E(\lambda_{t+1}a(1-\lambda_t))(z_{t+1})) z_{t+1}^{\frac{-1}{\sigma}}.$$

Solving for λ_t and rearranging:

$$\lambda_t = \frac{\gamma_t z_t^{\frac{1-\sigma}{\sigma}} \left(E(\lambda_{t+1})(z_{t+1})^{\frac{\sigma-1}{\sigma}}\right)}{1 + \gamma z_t^{\frac{1-\sigma}{\sigma}} \left(E(\lambda_{t+1})(z_{t+1})^{\frac{\sigma-1}{\sigma}}\right)},$$

where $\gamma = \left(\beta^{\frac{-1}{\sigma}} a^{1-\frac{1}{\sigma}}\right)$; and hence

$$\lambda_t = \frac{1}{1 + z_t^{\frac{\sigma-1}{\sigma}} \gamma^{-1} \left(E(\lambda_{t+1})(z_{t+1})^{\frac{\sigma-1}{\sigma}}\right)^{-1}} =$$

$$= \frac{1}{\left(1 + z_t^{\frac{\sigma-1}{\sigma}} \gamma^{-1} \left(E\left(\left(1 + z_{t+1}^{\frac{\sigma-1}{\sigma}} \gamma^{-1} \left(E(\lambda_{t+2})(z_{t+2})^{\frac{\sigma-1}{\sigma}}\right)^{-1}\right)^{-1}\right)(z_{t+1})^{\frac{\sigma-1}{\sigma}}\right)^{-1}\right)}$$

Redefine $\tilde{z}_t = z_t^{\frac{\sigma-1}{\sigma}}$. We then guess for a solution of the form:

$$
\begin{aligned}
\lambda_t &= \frac{1}{1 + E\sum_{s=t}\Pi_t^s \tilde{z}_s \left(\tilde{z}_{s+1}\right)^{-1}\gamma^{t-s-1}} \\
&= \frac{1}{1 + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s \left(\tilde{z}_{s+1}\right)^{-1}\gamma^{t+1-s-1}} \\
&= \left(1 + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s \left(\tilde{z}_{s+1}\right)^{-1}\gamma^{t+1-s-1}\right)^{-1} \\
&= \left(1 + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} + \gamma^{-1}\tilde{z}_t E\left(\tilde{z}_{t+1}\right)^{-1} E\sum_{s=t+1}\Pi_{t+1}^s \gamma^{t+1-s-1}\right)^{-1}
\end{aligned}
$$

If the guess is correct,

$$
E\lambda_{t+1}\tilde{z}_{t+1} = E\left(1 + E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s \left(\tilde{z}_{s+1}\right)^{-1}\gamma^{t+1-s-1}\right)^{-1}\tilde{z}_{t+1}
$$

Substitute the guess into λ_t to check:

$$
\lambda_t = \frac{1}{1 + \tilde{z}\gamma^{-1}\left(E\lambda_{t+1}\tilde{z}_{t+1}\right)^{-1}},
$$

and hence

$$
\begin{aligned}
\lambda_t &= \frac{1}{1 + \tilde{z}_t\gamma^{-1}E\left(1 + E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s (\tilde{z}_{s+1})^{-1}\gamma^{t+1-s-1}\right)(\tilde{z}_{t+1})^{-1}} \\
&= \frac{1}{1 + \tilde{z}_t\gamma^{-1}E(\tilde{z}_{t+1})^{-1}\left(1 + E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s (\tilde{z}_{s+1})^{-1}\gamma^{t+1-s-1}\right)}
\end{aligned}
\tag{18}
$$

Rearranging,

$$
\begin{aligned}
\lambda_t &= \left(1 + \tilde{z}_t\gamma^{-1}E\left(\tilde{z}_{t+1}\right)^{-1} + \tilde{z}_t\gamma^{-1}E\left(\tilde{z}_{t+1}\right)^{-1} E\sum_{s=t+1}\Pi_{t+1}^s \tilde{z}_s \left(\tilde{z}_{s+1}\right)^{-1}\gamma^{t+1-s-1}\right)^{-1} \\
&= \left(1 + \tilde{z}_t\gamma^{-1}E\left(\tilde{z}_{t+1}\right)^{-1} + \tilde{z}_t\gamma^{-1}E\left(\tilde{z}_{t+1}\right)^{-1} E\sum_{s=t+1}\Pi_{t+1}^s \gamma^{t+1-s-1}\right)^{-1}
\end{aligned}
$$

We conclude that the guess is in fact correct. ∎

Given an exogenous process $\lambda_t = \lambda(z_t)$, let

$$
V_\lambda\left(k_t\right) = (1-\sigma)^{-1}\left(\lambda(z_t)ak_t\right)^{1-\sigma} + E\left[\sum_{\tau=t+1}^{\infty}\beta^{\tau-t}\left(\lambda(z_\tau)ak_\tau\right)^{1-\sigma}\right]
$$

27

It follows that $V_\lambda(k_t)$ can be written as

$$V_\lambda(k_t) = m_t^\lambda(k_t)^{1-\sigma},$$

where

$$
\begin{aligned}
m_t^\lambda &= (1-\sigma)^{-1}(\lambda_t a)^{1-\sigma} + E\left((1-\sigma)^{-1}(\lambda_{t+1}a)^{1-\sigma}\right)\left[\beta\left((1-\lambda_t)a\right)^{1-\sigma}\right] \\
&\quad + E\sum_{s=t+2}^\infty \left((1-\sigma)^{-1}(\lambda_s a)^{1-\sigma}\right)\left[\prod_{j=t+1}^{s-1}\beta\left((1-\lambda_j)a\right)^{1-\sigma}\right] \\
&= (1-\sigma)^{-1}(\lambda_t a)^{1-\sigma} + \left((1-\sigma)^{-1}(a)^{1-\sigma}\right)\left[\beta\left((1-\lambda_t)a\right)^{1-\sigma}\right]E(\lambda_{t+1})^{1-\sigma} \\
&\quad + E\sum_{s=t+2}^\infty \left((1-\sigma)^{-1}(\lambda_s a)^{1-\sigma}\right)\left[\prod_{j=t+1}^{s-1}\beta\left((1-\lambda_j)a\right)^{1-\sigma}\right]
\end{aligned}
$$

Consider then the following maximization problem at time t, given Em_{t+1}^λ:

$$\max_{\lambda_t} \frac{(\lambda_t z_t ak_t)^{1-\sigma}}{(1-\sigma)} + \beta Em_{t+1}^\lambda\left((1-\lambda_t)ak_t\right)^{1-\sigma} \tag{19}$$

Lemma A.2: The solution to the maximization problem (19), λ_t, is *i)* increasing in z_t, and *ii)* decreasing in Em_{t+1}^λ.
Proof. The first order conditions for the maximization include:

$$(\lambda_t z_t ak_t)^{-\sigma} z_t ak_t = (1-\sigma)^{-1}\beta Em_{t+1}^\lambda\left((1-\lambda_t)ak_t\right)^{-\sigma}ak_t$$

which can be written as:

$$\lambda_t = \left(1 + \left((z_t)^{\sigma-1}(1-\sigma)^{-1}\beta Em_{t+1}^\lambda\right)^{\frac{1}{\sigma}}\right)^{-1}. \tag{20}$$

As a consequence, from (20),

$$\frac{d\lambda_t}{dz_t} > 0 \tag{21}$$

$$\frac{d\lambda_t}{dEm_{t+1}^\lambda} < 0 \tag{22}$$

that is, $\lambda(z_t)$ is increasing in z_t; and $\lambda(z_t)$ decreases in Em_{t+1}^λ. ∎

Proposition 1: Proof. Write the maximization problem (10) and the Markov Perfect Nash equilibrium problem (12-13) in the form of problem (19). Let Em_{t+1}^* and Em_{t+1}^M denote, respectively, the expected future value of the program evaluated at the solution of (10) and at (12-13). Note that $Em_{t+1}^M < Em_{t+1}^*$, since $\lambda^*(z_t)$ by definition maximizes

28

Em_{t+1}^{λ} with respect to λ. But then, (22) implies that $\lambda^M(z_t) > \lambda^*$, for any z_t. ■

Proposition 2: Proof. In the context of this proof, since we assume that $a_t = a > 0$, we can drop without loss of generality the state variable a_t from the notation. Existence of the value function $D(k_t, z_t)$ follows by Blackwell's Theorem by a standard argument. Moreover, it is straightforward to show that $D(k_t, z_t)$ is increasing in k_t. Let the policy function be denoted $\lambda(k_t, z_t)$.

Let

$$\lambda^E(k_t) = \arg\max_{\lambda} U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right]$$

Let $\lambda^{II}(k_t, z_t) = \max\{\lambda^E(k_t), \lambda^I(k_t, z_t)\}$. Then (5), that is,

$$D(k_t, z_t) =$$
$$\max \left[\begin{array}{l} \max_{\lambda \geq \lambda_t^I} U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right], \\ \max_{\lambda} U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right] - b(a k_t)^{1-\sigma} \\ \text{with } \lambda_t^I = \lambda^I(k_t, z_t) \end{array} \right]$$

can be written as

$$D(k_t, z_t) =$$
$$\max \left[\begin{array}{l} U(\lambda_t^{II} a k_t) + \beta E\left[D\left(\left(1-\lambda_t^{II}\right) a k_t, z_{t+1}\right)\right], \\ \max_{\lambda} U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right] - b(a k_t)^{1-\sigma} \\ \text{with } \lambda_t^{II} = \lambda^{II}(k_t, z_t) \end{array} \right]$$

We will now show that the policy function satisfies a cut-off rule, that is:

$$\lambda(k_t, z_t) =$$

$$\begin{cases} \lambda^{II}(k_t, z_t) & \text{for} \quad \lambda^I(k_t, z_t) \leq \overline{\lambda}(k_t) \\ \lambda^E(k_t) & \text{else} \end{cases}$$

We will then show that the cut-off, hence the policy function, are independent of k_t. Finally, we will prove the statement $\lambda^E > \lambda^*$.

The cut-off rule follows if we can show concavity of $U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right]$ with respect to λ. Fix k_t. Concavity guarantees that

$$\max_{\lambda} U(\lambda a k_t) + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right]$$

has a unique solution, λ^E, independent of the realization z_t. It follows that

$$(1-\sigma)^{-1}\left(\lambda^E a k_t\right)^{1-\sigma} + \beta E\left[D\left(\left(1-\lambda^E\right) a k_t, z_{t+1}\right)\right] - b(a k_t)^{1-\sigma} =$$

$$(1-\sigma)^{-1}\left(\lambda a k_t\right)^{1-\sigma} + \beta E\left[D\left((1-\lambda)\, a k_t, z_{t+1}\right)\right]$$

29

is satisfied for a value of λ, $\overline{\lambda} > \lambda^E$. By construction,

$$\frac{\partial}{\partial \lambda} \left[(1 - \sigma)^{-1} (\lambda a k_t)^{1-\sigma} + \beta E \left[D \left((1 - \lambda) a k_t, z_{t+1} \right) \right] \right] \leq 0 \text{ at } \lambda = \overline{\lambda}$$

and $\overline{\lambda}$ represents the cut-off for given k_t. Since k_t is arbitrary in the argument, we can construct in fact the cut-off $\overline{\lambda}(k_t)$ of the statement.

We turn now to show the concavity of

$$U (\lambda a k_t) + \beta E \left[D \left((1 - \lambda) a k_t, z_{t+1} \right) \right]$$

with respect to λ. It requires $U'' a k_t + \beta E \left[a k_t \frac{\partial^2}{\partial (k_{t+1})^2} D(k_t, z_t) \right] < 0$, and hence, in turn, $\frac{\partial^2}{\partial (k_{t+1})^2} D(k_t, z_t) < 0$.

Let $q_t = a k_t$. Choose arbitrary concave functions $h, U : R_+ \times R_+ \to R_+$ where $R_+ = [0, \infty)$, that is h, U take non-negative values. In particular, we can choose $U = (1 - \sigma)^{-1} c^{(1-\sigma)}$, $0 < \sigma < 1$. Let the operator T defined as follows:

$$(Th) (q_t; z_t) = \max \left[\begin{array}{c} U(\lambda_t^{II}(z_t) q_t) + \beta E \left[h \left(\left(1 - \lambda_t^{II}(z_t) \right) q_t, z_{t+1} \right) \right], \\ \max_\lambda U (\lambda q_t) + \beta E \left[h \left((1 - \lambda) q_t, z_{t+1} \right) \right] - b(q_t)^{1-\sigma} \end{array} \right] \qquad (23)$$

To show that $D(k_t, z_t)$ is concave, it suffices to show that the operator T preserves the concavity of the map h. Let $q = v q_t^1 + (1 - v) q_t^2$. From concavity of U and h, it follows that:

$$(Th) (q_t; z_t) \geq \max \left[\begin{array}{c} v \left[U(\lambda_t^{II}(z_t) q_t^1) + \beta E \left[h \left(\left(1 - \lambda_t^{II}(z_t) \right) q_t^1, z_{t+1} \right) \right] \right] \\ + (1 - v) \left[U(\lambda_t^{II}(z_t) q_t^2) + \beta E \left[h \left(\left(1 - \lambda_t^{II}(z_t) \right) q_t^2, z_{t+1} \right) \right] \right], \\ v \left[\max_\lambda U (\lambda q_t^1) + \beta E \left[h \left((1 - \lambda) q_t^1, z_{t+1} \right) \right] - b(q_t)^{1-\sigma} \right] \\ + (1 - v) \left[\max_\lambda U (\lambda q_t^2) + \beta E \left[h \left((1 - \lambda) q_t^2, z_{t+1} \right) \right] - b(q_t)^{1-\sigma} \right] \end{array} \right]$$

$$\geq \max \left[\begin{array}{c} \max \left[\begin{array}{c} v U(\lambda_t^{II}(z_t) q_t^1) + \beta E \left[h \left(\left(1 - \lambda_t^{II}(z_t) \right) q_t^1, z_{t+1} \right) \right], \\ v \left[\max_\lambda U (\lambda q_t^1) + \beta E \left[h \left((1 - \lambda) q_t^1, z_{t+1} \right) \right] - b(q_t)^{1-\sigma} \right] \end{array} \right], \\ \max \left[\begin{array}{c} (1 - v) U(\lambda_t^{II}(z_t) q_t^2) + \beta E \left[h \left(\left(1 - \lambda_t^{II}(z_t) \right) q_t^2, z_{t+1} \right) \right], \\ (1 - v) \left[\max_\lambda U (\lambda q_t^2) + \beta E \left[h \left((1 - \lambda) q_t^2, z_{t+1} \right) \right] - b(q_t)^{1-\sigma} \right] \end{array} \right] \end{array} \right]$$

The latter follows from $\max (a + b, c + d) \geq \max (a, c, b, d) = \max (\max (a, c), \max (b, d)) \geq 0$ if $a, b, c, d \geq 0$. Therefore,

$$(Th) (q; z_t) \geq \left[v (Th) (q_t^1; z_t) \right] + (1 - v) (Th) (q_t^2; z_t) \qquad (24)$$

and $(Th) (q_t; z_t)$ is concave.

We turn now to the independence of the policy function from k_t. The cut-off $\overline{\lambda}(a, k_t)$ solves equation

$$\max_\lambda U (\lambda a k_t) + \beta E \left[D \left(a_{t+1}, a_{t+1} (1 - \lambda) a k_t, z_t \right) \right] - b(a k_t)^{1-\sigma}$$
$$= U (\lambda a k_t) + \beta E \left[D \left((1 - \lambda) a k_t, z_t \right) \right]$$

in λ. Consider

$$D(k_t, z_t) = \max \left[\begin{array}{c} U(\lambda_t^{II} ak_t) + \beta E \left[D \left(\left(1 - \lambda_t^{II}\right) ak_t, z_{t+1} \right) \right], \\ \max_\lambda U \left(\lambda ak_t \right) + \beta E \left[D \left((1 - \lambda) ak_t, z_{t+1} \right) \right] - b(ak_t)^{1-\sigma} \end{array} \right]$$

Guess the following functional form for $D(k_t, z_t)$:

$$D(k_t, z_t) = M(z_t) (ak_t)^{1-\sigma}$$

Then,

$$M(z_t) (ak_t)^{1-\sigma} = \max \left[\begin{array}{c} (\lambda_t^{II} ak_t)^{1-\sigma} + \beta EM(z_{t+1}) \left(\left(1 - \lambda_t^{II}\right) ak_t \right)^{1-\sigma} \\ \max_\lambda (\lambda ak_t)^{1-\sigma} + \beta EM(z_{t+1}) \left((1 - \lambda) ak_t \right)^{1-\sigma} - b(ak_t)^{1-\sigma} \end{array} \right]$$

$$M(z_t) (ak_t)^{1-\sigma} = \max \left[\begin{array}{c} (\lambda_t^{II})^{1-\sigma} + \beta EM(z_{t+1}) \left(\left(1 - \lambda_t^{II}\right) \right)^{1-\sigma} \\ \max_\lambda (\lambda)^{1-\sigma} + \beta EM(z_{t+1}) \left((1 - \lambda) \right)^{1-\sigma} - b \end{array} \right] (ak_t)^{1-\sigma}$$

$$M(z_t) = \max \left[\begin{array}{c} (\lambda_t^{II})^{1-\sigma} + \beta EM(z_{t+1}) \left(\left(1 - \lambda_t^{II}\right) \right)^{1-\sigma} \\ \max_\lambda (\lambda)^{1-\sigma} + \beta EM(z_{t+1}) \left(a(1 - \lambda) \right)^{1-\sigma} - b \end{array} \right] \qquad (25)$$

It follows that the policy function $\lambda(z_t)$ associated with the dynamic program (25) is also the policy function associated with the program (5), and hence is independent of k_t. Furthermore, then, the cut-off is also independent of k_t: $\overline{\lambda}(k_t) = \overline{\lambda}$.

It remains to prove the statement $\lambda^E > \lambda^*$. Note that

$$\lambda^E = \arg \max_\lambda \lambda^{1-\sigma} + \beta EM(z_{t+1}) \left((1 - \lambda) \right)^{1-\sigma} \qquad (26)$$

The first order conditions of this maximization problem readily imply that λ^E decreases with an increase of $E[M(z_{t+1})]$. Moreover, it is straightforward to show that $E[M(z_{t+1})]$ decreases with b. But λ^* equals λ^E for $b = 0$. We conclude that, for any $b > 0$, $\lambda^E > \lambda^*$. $\blacksquare$

The **proof of Proposition 3** follows as an immediate corollary of Proposition 2, using the assumption that $\lambda^I(z)$ is increasing.

Proposition 4: Proof. Consider the maximization problems defining the two automatic processing mechanisms, (14) and (12-13), respectively, written in the form of problem (19). In the first case $Em_{t+1}^I = Em_{t+1}^*$ (under the incorrect belief that $z_\tau = 1$, $\tau \geq 1$); while in the second case $Em_{t+1} = Em_{t+1}^M$. We already noticed in the proof of Proposition 1 that $Em_{t+1}^* > Em_{t+1}^M$. We therefore conclude, by Lemma A.2, that $\lambda^M(z_t)$ is greater than $\lambda^I(z_t)$, when $\lambda^I(z_t)$ is determined by (14).

We next prove the statements in *i)* and *ii).* *i)* follows simply by continuity, since $\lambda^I(1) = \lambda^*$ (when $\lambda^I(z_t)$ is determined by (14)). To prove *ii)* instead notice that $EM(z_{t+1}) > Em^M_{t+1}$, since $M(z_t)$ is maximal for controlled processing and the Markov Perfect Nash equilibrium consumption-saving rule is feasible. When automatic processing is determined by (12-13), the statement then follows from Lemma A.2. ■

Proposition 5: Proof. By Proposition 8, $\lambda^M(z_t)$ is greater than $\lambda^I(z_t)$, when $\lambda^I(z_t)$ is determined by (14). The expected future value of the cognitive control program $EM(z_{t+1})$ is therefore larger when automatic processing is determined by (14). This is because the value when automatic processing is determined by (12-13) is feasible (but not maximal) when automatic processing is determined by (14). The result now follows from noticing that, as we have shown in the proof of Proposition 2, that λ^E decreases with $EM(z_{t+1})$. ■

Proposition 6: Proof. The proof is a straightforward corollary of Proposition 4 and 5. By Proposition 4, in fact $\lambda^E < \lambda^M(z_t)$, for any z_t, when λ^E is associated to automatic processing determined by (12-13). Moreover, by Proposition 5, λ^E is smaller when associated to automatic processing determined by (14) rather than (12-13). In this case also, therefore, $\lambda^E < \lambda^M(z_t)$. ■

Proposition 7: Proof. Write the Markov Perfect Nash equilibrium problem (12-13) in the form of problem (19). It is immediate to see that Em^M_{t+1} is decreasing in a first order stochastic dominance increase in the distribution of z_τ, $\tau > t$. But then, (22) implies that $\lambda^M(z_t)$ increases, for any z_t.

We study next the dependence of λ^E on first order stochastic dominance changes in the distribution of z_τ, $\tau > t$. We keep $\lambda^I(z_t)$ fixed in the argument. This is the case if automatic processing is determined by (14). We leave to the reader will check that the proof generalizes if $\lambda^I(z_t)$ increases with a first order stochastic dominance increase in the distribution of z_τ, $\tau > t$; which the case if automatic processing is determined by (12-13).

Consider dynamic program (25) that, we have shown in the proof of Proposition 2, characterizes $\lambda(z_t)$:

$$M(z_t) = \max \left[\begin{array}{c} (\lambda^{II}_t)^{1-\sigma} + \beta EM(z_{t+1})\left(\left(1 - \lambda^I_t\right)\right)^{1-\sigma} \\ \max_\lambda \lambda^{1-\sigma} + \beta EM(z_{t+1})\left((1 - \lambda)\right)^{1-\sigma} - b \end{array} \right] \qquad (27)$$

where $\lambda^{II}_t = \max\{\lambda^E, \lambda^I_t\}$.

The characterization of the cut-off rule in Proposition 2 implies that $M(z_t)$ is independent of z_t, for $z_t > \overline{z}$. Moreover $M(z_t)$ is decreasing in z_t, for $z_t \leq \overline{z}$ and such that $\lambda^I(z) > \lambda^E$. This is because

$$\lambda^{1-\sigma} + \beta EM(z_{t+1})\left((1 - \lambda)\right)^{1-\sigma}$$

32

is concave in λ.

Consider a first order stochastic dominance increase in the distribution of z_t. Such a change has then the effect of decreasing $EM(z_t)$; an effect which cannot be undone by a change in the cut-off without contradicting the definition of $M(z)$ as a value function, equation (27).

We pass now on to analyze the following problem

$$\arg\max_{\lambda}(\lambda)^{1-\sigma} + \beta EM(z_{t+1})((1-\lambda))^{1-\sigma} \tag{28}$$

which, by Proposition 2 is equivalent to problem

$$\arg\max_{\lambda} U(\lambda ak_t) + \beta E[D((1-\lambda)ak_t, z_{t+1})]$$

which appears in the statement.

The first order conditions of this maximization problem readily imply that λ increases with a decrease of $EM(z_{t+1})$, that is with a first order stochastic dominance increase in the distribution of z_t.

We study next the dependence of $\overline{\lambda}$ on first order stochastic dominance changes in the distribution of z_τ, $\tau > t$. Let $F(z_t)$ denote the cumulative distribution of z_t. Take a distribution $G(z_t)$ which dominates $F(z_t)$ in the first order stochastic sense, and consider the distribution obtained by mixing $F(z_t)$ with $G(z_t)$:

$$H(z_t) = (1-\alpha)F(z_t) + \alpha G(z_t)$$

Recall that, by an infinitesimal increase in the first order dominance sense in the distribution of z_t we mean an infinitesimal increase $d\alpha > 0$ at $\alpha = 0$.

Given b and $EM(z_{t+1})$, the cut-off $\overline{\lambda}$ is a solution of the following equation:

$$(\overline{\lambda})^{1-\sigma} + \beta EM(z_{t+1})\left(1-\overline{\lambda}\right)^{1-\sigma} = (\lambda^E)^{1-\sigma} + \beta EM(z_{t+1})\left(1-\lambda^E\right)^{1-\sigma} - b \tag{29}$$

where $\lambda^E = \arg\max_{\lambda} \lambda^{1-\sigma} + \beta EM(z_{t+1})(1-\lambda)^{1-\sigma}$.

Since $M(z_{t+1})$ is a continuous function, $d\alpha > 0$ has an infinitesimal negative effect on $EM(z_{t+1})$, that is $dEM(z_{t+1}) < 0$.

Given b and $EM(z_{t+1})$ the cut-off $\overline{\lambda}$ is determined by equation (29), where $\lambda^E = \arg\max_{\lambda} \lambda^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}(1-\lambda)^{1-\sigma}$. By the Envelope theorem, $(\lambda^E)^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}\left(1-\lambda^E\right)^{1-\sigma}$ is unaffected by any infinitesimal change $dEM(z_{t+1})$.

Once again, since $\overline{\lambda} > \lambda^E$ by construction of the cut-off, in Proposition 2, and $\lambda^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}(1-\lambda)^{1-\sigma}$ is concave in λ, it follows that $\lambda^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}(1-\lambda)^{1-\sigma}$ is in fact decreasing in λ at $\lambda = \overline{\lambda}$. The Implicit Function Theorem on (29) now implies that $\overline{\lambda}$ is locally decreasing in $EM(z_{t+1})$. ∎

Proposition 8: Proof. Note first that $\lambda^I(z_t)$ is independent of b, both if automatic processing is determined by (14) or by (12-13).

We study first the dependence of λ^E on an increase in b. Such a change has the straightforward effect of decreasing $EM(z_t)$. The first order conditions of (28) then readily imply that λ increases with a decrease of $EM(z_{t+1})$, that is with an increase in b.

We pass now to the analysis of the dependence of $\overline{\lambda}$ on an increase in b. Given b and $EM(z_{t+1})$, the cut-off $\overline{\lambda}$ is a solution of equation (29), where $\lambda^E = \arg\max_\lambda \lambda^{1-\sigma} + \beta EM(z_{t+1})(1-\lambda)^{1-\sigma}$ depends on b only through $EM(z_{t+1})$. From the definition of $M(z_t)$, equation (27), it follows in a straightforward manner that $EM(z_{t+1})$ is decreasing in b. Finally, since $\overline{\lambda} > \lambda^E$ by construction of the cut-off, in Proposition 2, and $(\lambda)^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}(1-\lambda)^{1-\sigma}$ is concave in λ, it follows that $(\lambda)^{1-\sigma} + \beta EM(z_{t+1}) E(a)^{1-\sigma}(1-\lambda)^{1-\sigma}$ is in fact decreasing in λ at $\lambda = \overline{\lambda}$. The Implicit Function theorem on (29) now implies that $\overline{\lambda}$ is locally increasing in b. ∎

Proof. ∎

We leave to the reader the straightforward **proof of Proposition 9**.

References

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

G. Ainsle (2001): *Breakdown of Will*, Cambridge University Press.

G. Ainsle and N. Haslam (1992): 'Hyperbolic Discounting,' in G. Lowenstein and J. Elster (eds.), *Choice Over Time*, Russell Sage Foundation.

J. Ameriks, A. Caplin, and J. Leahy (2002): 'Wealth Accumulation and the Propensity to Plan,' mimeo, NYU.

Y. Amihud and H. Mendelson (1986): "Asset Pricing and the Bid-Ask Spread,' *Journal of Financial Economics*, 17, 223-49.

G.M. Angeletos, D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001): 'The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation,' *Journal of Economic Perspectives*, 15, 3, 47-68.

R. A. Barkley (1997): *ADHD and the Nature of Self-Control*, Guilford Press.

Baumeister, R., T.F. Heatherton, and D.M. Tice (1994): *Losing Control: How and Why People Fail at Self Regulation*, San Diego, Academic Press.

A. Bechara, A. Damasio, H. Damasio, and S. Anderson (1994): 'Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex,' *Cognition*, 7-15.

A. Bechara, D. Tranel, H. Damasio, and A.R. Damasio (1996): 'Failure to Respond Autonomically to Anticipated Future Outcomes Following Damage to Prefrontal Cortex,' *Cerebral Cortex*, 6, 215-25.

R. Benabou and J. Tirole (2000): 'Willpower and Personal Rules,' mimeo, Princeton University.

D. Bernheim, and A. Rangel (2001): 'Addiction, Cognition and the Visceral Brain,' mimeo, Stanford University.

T. Braver and J. Cohen (2000): On the Control of Control: The Role of Dopamine in Regulating Prefrontal Function and Working Memory, MIT Press.

T. Braver, J.D. Cohen, and D. Servan Schreiber (1995): 'A Computational Model of Prefrontal Cortex Function,' in D.S. Touretzky, G. Tesauro, and T.K. Leen (Eds.), *Advances in Neural Information Processing Systems*, MIT Press.

M.J. Brennan, T. Chordia, A. Subrahmanyam (1998): 'Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns,' *Journal of Financial Economics*, 49, 345-73.

M. Browning and T.F. Crossley (2001): 'Unemployment Benefit Levels and Consumption Changes,' *Journal of Public Economics*, 80, 1, 1-23.

M. Browning and A. Lusardi (1996): 'Household Saving: Micro Theories and Micro Facts,' *Journal of Economic Literature*, 34, 1797-855.

J.J. Choi, D. Laibson, B.C. Madrian, and A. Metrick (2003): 'Consumption-Wealth Comovement of the Wrong Sign,' Harvard University, mimeo.

J.D. Cohen, K. Dunbar, and J.L. McClelland (1990): 'On the Control of Automatic Processes: A Parallel Distributed Processing Model of the Stroop Effect,' *Philosophical Transactions of the Royal Society (London) B*, 351, 1515-27.

J.D. Cohen, W.M. Perlstein, T.S. Braver, L.E. Nystrom, and D.C. Noll (1997): 'Temporal Dynamics of Brain Activation during a Working Memory Task,' *Nature*, 386, 604-8.

de Villiers, P.A. and R.J. Herrnstein (1976): 'Towards a Law of Response Strength,' *Psychological Bulletin*, 83, 1131-53.

J. Elster (1979): *Ulysses and the Sirens: Studies in Rationality and Irrationality*, Cambridge University Press, Cambridge.

R.W. Engel, M. Kane, and S. Tuholski (1999): 'Individual Differences in Working Memory Capacity and What They Tell Us About Controlled Attention, General Fluid Intelligence, and Functions of the Prefrontal Cortex,' in A. Miyake and P. Shah (Eds.), *Models of Working Memory: Mechanism of Active Maintenance and Executive Control*, Cambridge University Press.

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

S. Freud (1927) *The Ego and the Id*, Hogarth.

M. Friedman (1956): *A Theory of the Consumption Function*, Princeton University Press.

J. Fuster (1989): *The Prefrontal Cortex*, Raven Press, New York.

P.M. Gollwitzer (1999): 'Implementation Intentions: Strong Effects of Simple Plans,' *American Psychologist*, 54, 493-503.

P.M. Gollwitzer and J.A. Bargh (1996): *The Psychology of Action*, Guilford Press.

F. Gul and W. Pesendorfer (2001): 'Temptations and Self-Control,' *Econometrica*, 69, 6, 1403-35.

Herrnstein, R.J. (1961): 'Relative and Absolute Strengths of Response as a Function of Frequency of Reinforcement,' *Journal of the Experimental Analysis of Behavior*, 4, 267-72.

R.J. Herrnstein (1997): *The Matching Law: Papers in Psychology and Economics*, edited by H. Rachlin and D. Laibson, Russell Sage Foundation.

M.A. Just and P.A. Carpenter (1992): 'A Capacity Theory of Comprehension: Individual Differences in Working Memory,' *Psychological Review*, 99, 122-49.

Kirby, K.N. (1997): 'Bidding on the Future: Evidence Against Normative Discounting of Delayed Rewards,' *Journal of Experimental Psychology: General*, 126, 54-70.

K.N. Kirby and R.J. Herrnstein (1995): 'Preference Reversals Due to Myopic Discounting of Delayed Reward,' *Psychological Science*, 6, 83-9.

N.R. Kocherlakota (2001): 'Looking for Evidence of Time-Inconsistent Preferences in Asset Market Data,' *Federal Reserve Bank of Minneapolis Quarterly Review*, 25, 3, 13-24.

M.E. Kreinin (1961): 'Windfall Income and Consumption: Additional Evidence,' *American Economic Review*, 51, 3, 388-90.

J. Kuhl and J. Beckmann (Eds., 1985), *Action Control: From Cognition to Behavior*, Springer Verlag.

D. Laibson (1996): 'Golden Eggs and Hyperbolic Discounting,' *Quarterly Journal of Economics*, CXII, 443-77.

M. Landsberger (1966): 'Windfall Income and Consumption: A Comment,' *American Economic Review*, 56, 534-39.

F. Lhermitte (1986): 'Patient Behavior in Complex and Social Situations: The "Environmentally Dependency Syndrome",' *Annals of Neurology*, 19, 335-43.

Loewenstein, G. (1996): 'Out of Control: Visceral Influences on Behavior,' *Organizational Behavior and Human Decision Processes*, 65, 272-92.

Loewenstein, G., T. O'Donoghue, and M. Rabin (2002): 'Projection Bias in Predicting Future Utility,' Carnegie Mellon, mimeo.

E. Majani, R. Erlason, and Y. Abu Mostafa (1989): 'The Introduction of Multiscale Temporal Structure,' in D.S. Touretzky (ed.), *Advances in Neural Information Processing Systems, I*, San Mateo, CA, Morgan Kaufmann.

J.M. Manchester and J.M. Poterba (1989): 'Second Mortgages and Household Savings,' *Regional Science and Urban Economics*.

J.L. McClelland, D.E. Rumelhart, eds. (1986): *Parallel Distributed Processing*, Cambridge, MA, MIT Press.

E.K. Miller and J. Cohen (2001): 'An Integrative Theory of Prefrontal Cortex Function,' *Annual Review of Neuroscience*, 24, 167-202.

A. Miyake and P. Shah (Eds., 1999): *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, Cambridge University Press.

F. Modigliani and R. Brumberg (1954): 'Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data,' in K.K. Kurihara (Ed.), *Post-Keynesian Economics*, Rutgers University Press.

S. Monsell and J. Driver (Eds., 2000): *Control of Cognitive Processes: Attention and Performance XVIII*, MIT Press.

E. D. O'Donoghue and M. Rabin (1999): 'Doing It Now or Doing It Later,' *American Economic Review*, 89, 103-24.

R.C. O'Reilly (1999): 'Six Principles for Biologically-Based Computational Models of Cortical Cognition,' appeared in, *Trends in Cognitive Sciences*, 2, 455-62, (1998).

R.C. O'Reilly and Y. Munakata (2000): *Computational Explorations in Cognitive Neuroscience*, MIT Press.

J.A. Parker (1999): 'The Reaction of Household Consumption to Predictable Changes in Social Security Taxes,' *American Economic Review*, 89, 4, 959-73.

L. Pastor and R.F. Stambaugh (2001): 'Liquidity Risk and Expected Stock Returns,' NBER Working Paper 8462.

J.M. Poterba, S.F. Venti and D.A. Wise (2001): 'The Transition to Personal Accounts and Increasing Retirement Wealth: Macro and Micro Evidence,' NBER Working Paper 8610.

V. Prabhakaran, K. Narayanan, Z. Zhao, J.D. Gabrieli (2000): 'Integration of Diverse Information in Working Memory within the Frontal Lobe,' *Neuroscience*, 3, 85-90.

A. Rubinstein (2000): 'Is It "Economics and Psychology" ?: The case of Hyperbolic Discounting,' Tel Aviv University, mimeo.

W. Schultz (1998): 'Predictive Reward Signal of Dopamine Neurons,' *Journal of Neurophysiology*, 80, 1-27.

W. Schultz, P. Apicella, R. Romo, and E. Scarnati (1995): 'Context-Dependent Activity in Primate Striatum Reflecting Past and Future Behavioral Events,' in J.C. Houk, J.L. Davis and D.G. Beiser (Eds.), *Models of Information Processing in the Basal Ganglia*, MIT Press.

W. Schultz, P. Apicella, T. Ljungberg (1993): 'Responses of Monkey Dopamine Neurons to Rewards and Conditional Stimuli during Successive Steps of Learning a Delayed Response Task,' *Journal of Neuroscience*, 13, 900-13.

W. Schultz, P. Dayan, and P.R. Montague (1997): 'A Neural Substrate of Prediction and Reward,' *Science*, 275, 1593.

M.D. Shapiro and J. Slemrod (2002): 'Consumers Response to Tax Rebates,' forthcoming in *American Economic Review*.

H.M. Shefrin and R.H. Thaler (1992): 'Saving and Mental Accounting,' in G. Loewenstein and J. Elster (Eds.), *Choices over Time*, Russell Sage Foundation.

B. Shiv, and A. Fedorikhin (1999): 'Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making,' *Journal of Consumer Research*, 26, 278-92.

N.S. Souleles (2002): 'Consumer Response to the Reagan Tax Cuts,' *Journal of Public Economics*, 85, 99-120.

N.S. Souleles (2000): 'College Tuition and Household Savings and Consumption,' *Journal of Public Economics*, 77, 185-207.

N.S. Souleles (1999): 'The Response of Household Consumption to Income Tax Refunds,' *American Economic Review*, 89, 4, 947-58.

R.H. Strotz (1956): 'Myopia and Inconsistency in Dynamic Utility Maximization,' *Review of Economic Studies*, 23, 166-80.

R.H. Thaler (1990): 'Anomalies: Savings, Fungibility, and Mental Accounts,' *Journal of Economic Perspectives*, 4, 1, 193-205.

R.H. Thaler and H.M. Shefrin (1981): 'An Economic Theory of Self Control,' *Journal of Political economy,* 89(2), 392-406.

P. Vendrell, C. Junque, J. Pujol, M.A. Durado, J. Molet, and J. Grafman (1995): 'The Role of the Prefrontal Regions in the Stroop Task,' *Neuropsychologia*, 33, 341-52.

S.F. Venti and D.A. Wise (1987a): 'Have IRAs Increased U.S. Saving?: Evidence from Consumer Expenditure Surveys,' NBER Working Paper 2217.

S.F. Venti and D.A. Wise (1987b): 'But They Don't Want to Reduce Housing Equity,' NBER Working Paper 2859.

K.D. Vohs, and T.F. Heatherton (2000): 'Self-regulatory failure: A resource-depletion approach,' *Psychological Science*, 11, 249-254.

M. Warshawsky (1987): 'Sensitivity to Market Incentives: The Case of Policy Loans,' *Review of Economic and Statistics*, 286-95.

D.W. Wilcox (1989): 'Social Security Benefits, Consumption Expenditure, and the Life Cycle Hypothesis,' *Journal of Political Economy*, 97, 288-304.