Strategic Ignorance as a Self-Disciplining Device

JUAN D. CARRILLO
ECARES, Université Libre de Bruxelles and CEPR

and

THOMAS MARIOTTI
London School of Economics

We analyse the decision of an agent with time-inconsistent preferences to consume a good that exerts an externality on future welfare. The extent of the externality is initially unknown, but may be learned via a costless sampling procedure. We show that when the agent cannot commit to future consumption and learning decisions, incomplete learning may occur on a Markov perfect equilibrium path of the resulting intra-personal game. In such a case, each agent's incarnation stops learning for some values of the posterior distribution of beliefs and acts under self-restricted information. This conduct is interpreted as strategic ignorance. All equilibria featuring this property strictly Pareto dominate the complete learning equilibrium for any posterior distribution of beliefs.

1. INTRODUCTION

In a world of uncertainty, individual decisions are driven by perceptions of risks as well as preferences. While objective risk estimates seem the most relevant measure for decision making, there is considerable evidence that subjective and objective estimates are often far from each other. More importantly, *aggregate biases* in the perception of risks turn out to be pervasive. For example, Viscusi (1990) shows on the basis of a sample of 3119 individuals (including 779 smokers) that the average perceived probability of getting lung cancer because of smoking is 0·426 for the full sample and 0·368 for smokers. By contrast, the U.S. Surgeon General's estimate for this risk lies in a range from 0·05 to 0·10. In our view, this divergence between subjective and objective risk assessments can hardly be explained solely by inability or cost of acquiring information, since studies on the health effects of tobacco are widely publicized and often freely available.[1] An alternative explanation may be helpful to understand why individuals do not use or collect all available evidence.

In this paper, we argue that people may prefer to stay away from available information, fearing the impact that a change of belief could have on their behaviour. Non-smokers may for instance anticipate that optimistic estimates of tobacco's impact on health might induce them to smoke, with the risk of being trapped in overconsumption. This suggests that voluntary ignorance could be used as a self-control device preventing the individual from embarking on a hazardous activity which he may later regret. However, for ignorance to have such a commitment value, it is necessary to depart from the

---

1. In any case, information costs would not explain the systematic bias implied by the observation that as many as 94·8% of individuals in the full sample and 90·3% of the smokers overestimate the risk.

usual paradigm of a rational, time-consistent individual decision-maker. We shall rather consider the individual as a collection of incarnations with conflicting goals. Specifically, our theory is based on the following two building blocks.

First, *time inconsistency*. We focus on an individual with dynamically inconsistent preferences (Strotz (1956)). In each period, the instantaneous payoffs are overweighted relative to future rewards, so that the individual discounts short-term events at a higher rate than long-term events. At each date, a consumption decision can be made which raises instantaneous payoffs but exerts a negative externality on future welfare, just as smoking in the above discussion.[2] A crucial assumption is that the individual cannot commit to his future decisions, and therefore plays a non-cooperative game with his future incarnations.

Second, *costless learning and perfect recall*. There is incomplete information about a parameter that affects the magnitude (or frequency) of the externality. At every period, and before taking his consumption decision, the individual has the opportunity to collect information about this parameter at no cost, and to update his beliefs in a Bayesian way. If at some period the information acquisition process is exhaustive, complete knowledge of the parameter is achieved. Given perfect recall, all information gathered at some date can and will be used by the individual in the subsequent periods.

We focus on the intra-personal game played by the temporal incarnations (or "selves") of the individual, in which each incarnation decides first whether to collect or not information about the externality, and then whether to consume or not in the current period. We find that, even if learning is free and creates no delays for consumption, the individual may decide in equilibrium not to acquire all available information, a conduct that can be identified as *strategic ignorance*.

This result rests on the combination of informational and consumption externalities present in our model. Because of the dynamic inconsistency of preferences, optimal contingent plans from the perspective of a given self are no longer optimal for future selves. This, together with the inability to commit to a consumption path, may lock the individual into a state of systematic overconsumption. In that situation, ignorance may help to avoid such a time-inconsistent behaviour. Indeed, there is a trade-off in the decision to acquire information. On the one hand, under full information each incarnation can take the optimal consumption decision in the current period. On the other hand, this information is shared with all future incarnations. The latter take the optimal action *from their perspective*, which is suboptimal from the current viewpoint. Overall, each incarnation may prefer not to acquire information that could be used by future incarnations, at the cost of not being able to take the optimal current decision. If all incarnations follow the same reasoning, this behaviour is consistent with subgame-perfection.

The result is first illustrated in a simple three-period example with limited learning opportunities. Then, we consider a stationary infinite horizon model with unrestricted learning possibilities. In this general framework, we study the Markov perfect equilibria (MPE) of the intra-personal game, for which the distribution of posterior beliefs about the externality is the payoff relevant state variable. We show that our game always has (i) a *complete learning* MPE where the agent collects all available information before taking his first-period consumption decision, independently of his initial beliefs; and (ii) a family of *strategic ignorance* MPEs, each of them characterized by a set of distributions of posterior beliefs on which incarnations stop collecting information, and then choose whether

---

2. Alternatively, the current decision could be a costly effort which exerts a positive externality on future welfare.

to consume or not. Moreover, we prove that the welfare of *all incarnations* is strictly higher in any strategic ignorance MPE than in the complete learning MPE.

At this point, it may be helpful to be precise in which sense beliefs are manipulated in our model. Obviously, if the experimentation is not exhaustive, beliefs will not converge to the true value. Yet, since incarnations update their beliefs in a Bayesian way, each of them will not hold a biased belief, *on average*. This suggests that, in our model, only the higher-order moments of the belief distribution may and will be manipulated. However, *conditionally on the truth, there will be a systematic bias* on the expected value of the posterior about the externality. Indeed, in our binary consumption model, there is a risk threshold such that an individual consumes if and only if his estimate is below this threshold. We show that if the true risk happens to be below this threshold, then the *ex post* aggregate distribution will have (i) a mass of individuals who have learned the truth and consume accordingly, and (ii) a mass of individuals above this threshold who self-restrict their information gathering and refrain from consuming. Naturally, more comprehensive implications can be obtained when the set of alternatives is extended beyond the binary case. Overall, our model provides a robust and testable prediction: when consumption has current benefits and delayed costs then, conditionally on the true risk being "small", the *ex post* average perceived risk will be higher than the objective risk and the average consumption smaller than under full learning. This is precisely the empirical findings in Viscusi's (1990) study on the risks of tobacco.

Before proceeding with the formal analysis, several remarks are in order. First, we want to emphasize that the manipulation of information is endogenous, and of course depends on the realizations of the individuals' sampling. Second, the same logic about the value of strategic ignorance as a self-commitment device may be applied to procrastination problems. For example, a researcher with meager but encouraging information may undertake a project, while better information may cast some doubts about its quality and lead to beliefs involving "inefficient procrastination". Hence, our model has also a general prediction in activities subject to current costs and delayed benefits: conditionally on the truth, the *ex post* average perceived risk will be smaller than the objective risk and the average level of activity higher than under full learning. Last, note that our work can be seen as dual to the literature on learning by experimentation,[3] in the sense that learning increases *current* rewards but reduces *future* payoffs, as a consequence of future actions being suboptimal from the current perspective. In particular, there is no trade-off in our setting between high current rewards and accumulation of information leading to possible increases in future returns. Rather, what possibly restrains players from exhausting their learning opportunities is precisely the informational externality generated by the learning process.

2. CONSUMPTION WITH TIME-INCONSISTENT PREFERENCES

2.1. *The basic model*

The main elements of our model are the following:

• *Actor(s)*.   Time is discrete and indexed by $t = 0, 1, 2, \ldots$ We view the consumer as a countable collection of risk-neutral incarnations, with one incarnation per period. We call "self-t" the consumer's incarnation at date t.[4]

---

3. See, for example, Rothschild (1974), McLennan (1984), Easley and Kiefer (1988), Aghion, Bolton, Harris and Jullien (1991) or Keller and Rady (1999).

4. An alternative formulation provided in the literature to model conflict between selves considers a principal/multiagent structure where the "planner" (principal) maximizes the individual's intertemporal welfare, whereas each "doer" (agent) is solely concerned about current welfare (see Thaler and Shefrin (1981)).

• *Actions.* In every period, one unit of a free indivisible good is available for consumption. Let $x_t \in \{0, 1\}$ denote the amount consumed in period t.[5]

• *Externalities.* Consumption increases the instantaneous utility of the individual but exerts a negative externality on the welfare of his future incarnations. More specifically, we assume that a positive consumption level at any date t lowers the per-period payoffs of all subsequent selves $t + \tau$ ($\tau \geqq 1$) by an amount $\lambda^{\tau-1}C > 0$ with probability θ. Here, $\lambda \in [0, 1)$ is a depreciation factor which captures the fact that consumption at date t has a higher expected negative effect on nearer than on distant future incarnations. On the whole, the expected negative externality I_t imposed on self-t by his predecessors is given by

$$I_t = \sum_{\tau=0}^{t-1} \lambda^{t-\tau-1} x_\tau \theta C. \tag{2.1}$$

Note that since I_t is additively separable with respect to past consumption levels $\{x_\tau\}_{\tau=0}^{t-1}$, the marginal externality induced on future incarnations' welfare by one unit of consumption at the present date is *independent* of the past levels of consumption.

• *Information.* The probability of exerting the negative externality θ is unknown to the players. It is distributed according to some prior probability distribution π_0 with continuous density $f_0(\theta)$ bounded away from zero over the full support $[0, 1]$. However, each self can costlessly acquire information about θ and update his beliefs accordingly. We also assume that I_t is not observable at any date t. (The description of the information acquisition process is postponed until Section 3.2.)

• *Instantaneous payoffs.* The net instantaneous payoffs at each date t are linear

$$u_t = x_t - I_t. \tag{2.2}$$

As for the externality cost, the instantaneous marginal utility of consumption is independent of past consumption levels. Note that since I_t is not observable, the individual does not know his current payoff in any stage. Still, what matters for his decision-making is only the difference in utility between consuming and abstaining.

• *Intertemporal payoffs.* A key element of our theory is the existence of an *intrapersonal conflict among selves*. This conflict is captured by assuming that the consumer's preferences are dynamically inconsistent in the sense of Strotz (1956). We adopt the discount function introduced by Phelps and Pollak (1968). Specifically, from the perspective of self-t, the $t + \tau$-period ($\tau \geqq 1$) discount factor is set equal to $\beta \delta^\tau$, where $0 < \beta < 1$. The intertemporal utility index of self-t is then given by

$$U_t = E_t \left(u_t + \beta \sum_{\tau=1}^{\infty} \delta^\tau u_{t+\tau} \right). \tag{2.3}$$

The parameter β in (2.3) represents the "salience"[6] of current payoffs relative to the future stream of returns (which are, period to period, discounted at a rate δ). Equations (2.2) and (2.3) imply that the marginal rate of substitution between one unit of consumption at date t and one unit of consumption at date $t + 1$ is equal to $1/\beta\delta$ from the perspective of self-t, while it is equal to $1/\delta$ from the perspective of all previous selves. Hence,

---

5. In an earlier version of this paper (Carrillo and Mariotti (1998)), we show that the restriction to binary decisions is taken without loss of generality. See also Carrillo (1998) for a model of time inconsistent preferences and a continuous consumption decision.

6. This terminology is borrowed from Akerlof (1991). Psychologists would rather refer to β as an "impatience" or "impulsiveness" parameter, see Ainslie (1992).

optimal and contingent plans from the perspective of self-t are no longer optimal for subsequent incarnations, and are not enforceable should the consumer not be able to commit to a course of actions. Naturally, the closer β is to 1, the better self-t endogenizes the externality he exerts on future selves through his current consumption.

The above specification of intertemporal payoffs captures most of the properties of a generalized hyperbolic discount function $\mathscr{V}_{t,t'} = 1/(1 + b(t' - t))$, where t is the current period, $t' > t$ is the consumption period, and b is a constant. Equation (2.3) implies in particular that short-term events are discounted at a higher rate than long-term events. Since the pioneering work of Ainslie (1975), some scholars (*e.g.* Thaler (1981), Mazur (1987)) have emphasized the empirical relevance of hyperbolic discount functions for modelling both human and animal dynamic behaviour. We refer the reader to Ainslie (1992, Chapter 3), or Loewenstein and Prelec (1992) for a review of these contributions, and to Akerlof (1991), Laibson (1996, 1997) and Caillaud, Cohen and Jullien (1996) for applications to procrastination, consumption/savings decisions and addiction problems, respectively.

• *Foresight.* Last, we assume that the consumer perfectly anticipates his dynamically inconsistent behaviour and rationally behaves accordingly.[7]

As pointed out in the introduction, our model can be adapted to deal with any situation where decisions affect the welfare of current and future incarnations in opposite directions. Here we focus on the simplest case, in which the marginal cost and return of current decisions are independent of past behaviour. Introducing habit formation (*i.e.* a complementarity between the marginal utility of current consumption and the level of past consumption, as in Becker and Murphy (1988) for example) or risk-aversion in our model would only add a complementary dimension to the problem.

2.2. *A three-period example with limited learning opportunities*

Our first goal is to illustrate with a simple example the main result of the paper, namely the value of ignorance under time inconsistency. There are three periods, $t \in \{0, 1, 2\}$. The individual may either consume or abstain in periods 0 and 1, and learn the true value of θ only in period 0, before his consumption decision. Let $x_0, x_1 \in \{0, 1\}$ be the corresponding consumption levels. For simplicity, we assume that (i) $\delta = 1$, (ii) the externality C is exerted only in the period after consumption (*i.e.* $\lambda = 0$), and (iii) $1/\beta C < 1$. In this setup, the intertemporal utility from the perspective of each self is given by

$$U_0(x_0, x_1) = x_0(1 - \beta\theta C) + x_1\beta(1 - \theta C),$$

$$U_1(x_0, x_1) = -x_0\theta C + x_1(1 - \beta\theta C), \tag{2.4}$$

$$U_2(x_0, x_1) = -x_1\theta C.$$

According to $U_0(\cdot)$, self-0 would like to consume in both periods if $\theta \in [0, 1/C]$, to consume only in period 0 if $\theta \in (1/C, 1/\beta C)$ and to abstain in both periods if $\theta \in [1/\beta C, 1]$. However, he cannot commit on future decisions. Therefore, given $U_1(\cdot)$, one can see that if self-0 learns θ, the individual will end up consuming in both periods if $\theta < 1/\beta C$ and abstaining in both periods if $\theta \geq 1/\beta C$. Similarly, and given risk-neutrality, if self-0 does

7. The same assumption is made by Laibson (1996, 1997). By contrast, Akerlof (1991) considers an agent who fails to foresee his dynamic inconsistency. O'Donoghue and Rabin (1998, 1999) compare the welfare losses when agents have perfect foresight (in their terminology, "sophisticated") and imperfect foresight ("naive").

not learn θ the agent will consume in periods 0 and 1 if $E_{\pi_0}(\theta) < 1/\beta C$ and abstain in both if $E_{\pi_0}(\theta) \geq 1/\beta C$. Overall, self-0's expected payoff if he learns θ is given by

$$V_L = \pi_0(\theta < 1/\beta C)[1 + \beta - 2\beta E_{\pi_0}(\theta | \theta < 1/\beta C)C]. \tag{2.5}$$

By contrast, his expected payoff if he remains uninformed about θ is

$$V_{NL} = \begin{cases} 1 + \beta - 2\beta E_{\pi_0}(\theta)C & \text{if } E_{\pi_0}(\theta) < 1/\beta C, \\ 0 & \text{if } E_{\pi_0}(\theta) \geq 1/\beta C. \end{cases} \tag{2.6}$$

From (2.5) and (2.6) it is immediate that

(i) If $E_{\pi_0}(\theta) < 1/\beta C$, then $V_{NL} < V_L$.
(ii) If $E_{\pi_0}(\theta) \geq 1/\beta C$, then $V_{NL} > V_L$ if and only if $E_{\pi_0}(\theta | \theta < 1/\beta C) > (1 + \beta)/2\beta C$.

The intuition is simple. From $U_0(\cdot)$ and $U_1(\cdot)$ we note that the source of the intra-personal conflict is the existence of a set of values $\theta \in (1/C, 1/\beta C)$ such that self-0 would like to consume only in period 0 but ends up consuming in both periods. Therefore, a *necessary* condition for ignorance being valuable is that it induces abstention in period 1, which is the case when $E_{\pi_0}(\theta) \geq 1/\beta C$. However, this is not sufficient because ignorance also entails several costs. Indeed, from self-0's viewpoint, ignorance and abstention is suboptimal at date 0 if $\theta \in [0, 1/\beta C)$ and at date 1 if $\theta \in [0, 1/C]$. Note also that when $\theta \in [1/\beta C, 1]$, ignorance has neither costs nor benefits: the consumer (optimally) abstains both at dates 0 and 1. Overall, condition (ii) states that the benefits of ignorance offset the costs if, conditional on $\theta < 1/\beta C$, it is more likely that θ is close to $1/\beta C$ rather than close to 0.[8]

3. INFINITE HORIZON AND COSTLESS LEARNING

While the previous three-period case suggests that ignorance may have a positive value for a time-inconsistent decision maker, it is subject to several shortcomings. Its first limitation lies in the fact that, due to the assumption of a finite horizon, information has always a positive value for the last active incarnation of the consumer. In the above example, if self-1 had the opportunity to learn the true value of θ at no cost, he would always exert this option. Anticipating this impossibility to constrain the future choices, self-0 would therefore prefer to collect the information himself. More generally, in any finite horizon version of our model, if information is freely available at each date, perfect learning always occur at equilibrium. Second, in the example presented, remaining uninformed is optimal for self-0 only as long as condition (ii) is satisfied by the prior distribution of beliefs π_0. However, this limitation is due to our crude representation of the learning process, in which the opportunities faced by the consumer at date 0 are either full learning or no learning. It seems however that the logic of the argument should remain intact if we included the possibility of imperfect learning, by allowing the individual to sequentially collect pieces of information about θ.

In this section, we study an infinite horizon model with costless experimentation about θ in each period. Moreover, we consider a general learning technology in which the consumer can run sequential sampling experiments. This allows us to generalize the results of Section 2.2, and to provide new insights. Our first conclusion is that the learning

---

8. For instance, when θ is close to $1/\beta C$, ignorance has almost no cost $(1 - \beta\theta C \simeq 0)$ but some benefits $(\beta(1 - \theta C) < 0)$. Note that when $\beta = 1$ the condition stated in (ii) never holds: the intra-personal conflict is necessary for our result

strategy depends crucially on the time horizon. Indeed, we show that strategic ignorance is always an equilibrium of the infinite horizon game where information is freely available in each period. This conclusion holds independently of the consumer's initial belief, thus strengthening the case for strategic ignorance. In addition, an important new insight is that each self is now *strategically constrained* by the experimentation strategy of his successors. This additional conflict among selves can only be captured in a very crude way within a finite horizon model, since it precludes the very existence of a strategic ignorance equilibrium. A related point is that focusing on the Markov perfect equilibria of a stationary model allows an unobstructed analysis of the interactions between consumption and informational externalities that form the basis of our model. We show in particular that different levels of learning can be achieved depending on the degree of coordination among selves. An important consequence is that strategic ignorance leads to unambiguous Pareto improvements compared to complete learning.

3.1. *Benchmark case: no learning*

Consider first the benchmark situation where the consumer has no opportunity to acquire information at any date, so that his beliefs about θ remain at his prior π_0.[9] In the absence of learning, the agent's behaviour depends on his ability to commit to future consumption decisions. From the definition of intertemporal payoffs in (2.3), the impact on self-t's welfare of consumption x_t at date t is

$$x_t\left(1 - \beta\delta\sum_{s=1}^{\infty}(\lambda\delta)^{s-1}E_{\pi_0}(\theta)C\right) = x_t\left(1 - E_{\pi_0}(\theta)\frac{\beta\delta}{1-\lambda\delta}C\right). \tag{3.1}$$

Similarly, for each $\tau \geqq 1$, the impact on self-t's welfare of consumption $x_{t+\tau}$ at date $t+\tau$ is

$$\beta\delta^{\tau}x_{t+\tau}\left(1 - \delta\sum_{s=1}^{\infty}(\lambda\delta)^{s-1}E_{\pi_0}(\theta)C\right) = \beta\delta^{\tau}x_{t+\tau}\left(1 - E_{\pi_0}(\theta)\frac{\delta}{1-\lambda\delta}C\right). \tag{3.2}$$

Define $\theta^* = (1 - \lambda\delta)/\beta\delta C$. The following assumption will be maintained throughout the paper.

Assumption 1. $\theta^* < 1$.

Were this assumption violated, each incarnation would consume in any period as soon as he had discretion over current actions, whatever the behaviour of the other selves and the value of $E_{\pi_0}(\theta)$. We are in position to characterize the equilibrium of the game when the individual can commit at date 0 to all his future consumption decisions.

Lemma 1.   *Under full commitment at date 0 and no learning opportunities, the unique equilibrium is such that:*

   (i) *If $0 \leqq E_{\pi_0}(\theta) \leqq \beta\theta^*$, the agent consumes one unit per period;*
   (ii) *If $\beta\theta^* < E_{\pi_0}(\theta) < \theta^*$, the agent consumes one unit in period 0 and abstains thereafter;*
   (iii) *If $\theta^* \leqq E_{\pi_0}(\theta) \leqq 1$, the agent abstains from consuming in each period.*

Proof.   Immediate given (3.1) and (3.2). ||

9. Clearly, the results of this section extend to the case where the agent has complete information about θ.

The behaviour of the individual is modified if he cannot commit at any date to a continuation consumption path, so that self-t has discretion over decisions at stage t only. The following assumption will be maintained in the remainder of the paper.

Assumption 2.   No self can commit to the consumption decisions of his successors.

The reader might argue that some kinds of self-commitment are often feasible. Still, our objective is to highlight the role of information processing in impulse control. This is why we deliberately rule out the various kinds of commitment devices that have been proposed in the literature.[10]

Note that, due to the additive separability of intertemporal payoffs, the consumption path up to period $t-1$ is payoff irrelevant from self-t's perspective. In order to discard unnecessary dependence of equilibrium behaviour on past variables, we shall leave all the "bootstrap" equilibria out of the scope of the analysis and focus instead on Markov perfect equilibria (MPE) of the consumption game.

Lemma 2.   *Under no commitment and no learning opportunities, the unique* MPE *is such that*:

   (i) *If* $0 \leqq E_{\pi_0}(\theta) < \theta^*$, *each self consumes one unit whatever the amount consumed by other selves*;
   (ii) *If* $\theta^* \leqq E_{\pi_0}(\theta) \leqq 1$, *each self abstains whatever the amount consumed by other selves*.

Proof.   As for Lemma 1, the proof is immediate given (3.1) and (3.2). ||

It follows from Lemmas 1 and 2 that the no-commitment solution differs from the full-commitment solution only for the intermediary range of beliefs $(\beta\theta^*, \theta^*)$. By definition, for beliefs in this region, self-0 would be strictly better off constraining his future choices. More interestingly, for a subset of beliefs in that interval, self-0 is strictly worse off consuming at every period than if he could commit to abstain in each stage of the game, including the current one. Stated formally, denote $E_{\pi_0}(U_0(x; \theta))$ self-0's expected intertemporal utility if all selves consume $x \in \{0, 1\}$. From (2.1) and (2.3), it is easily shown that $E_{\pi_0}(U_0(1; \theta)) < E_{\pi_0}(U_0(0; \theta))$ as soon as $E_{\pi_0}(\theta) \in (\tilde{\theta}, \theta^*)$, where $\tilde{\theta} = [1 - \delta(1 - \beta)]\theta^*$. This interval will be referred to as the *inefficient consumption region* (IC). Moreover, the same reasoning is also true for all subsequent selves. In other words, for beliefs in IC, the consumer at each date is willing not only to constrain his future choices but *even to constrain his current choices provided that future selves act in the same way*. As we shall see, this feature of the MPE solution is crucial for our analysis when learning opportunities becomes available before each consumption decision.[11] This result is illustrated in Figure 1.

   10. This includes for instance private side bets (described by Ainslie (1992) and formalized in Caillaud *et al.* (1996)) or mental accounting (Thaler (1990)).
   11. Under more general hyperbolic discounting, it may be optimal for self-0 to consume during more than one period before stopping. However, our results are robust in the following sense: (i) under commitment, once the agent gives up consuming at the optimum, his decision is irrevocable and (ii) under no commitment there always exists a region such that no consumption even in the current period dominates consumption at every date
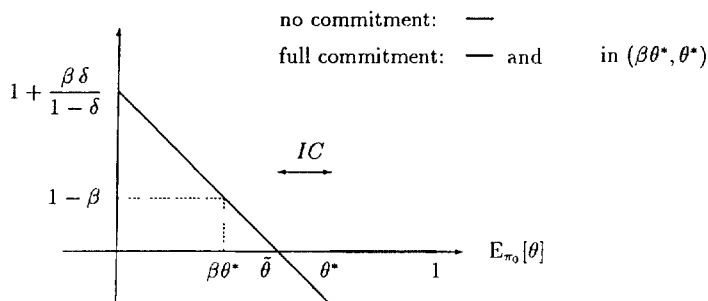
FIGURE 1

Self-0's expected intertemporal utility in the full commitment and the no commitment cases

3.2. *Learning and strategies*

The learning technology can be described as follows:

- *Information acquisition.*   Learning is modelled as a Bernoulli sampling process. Before taking his consumption decision, each self has the opportunity to gather an arbitrary (and possibly) infinite number of pieces of information about θ. Each sample z is a signal either of the innocuousness of consumption (with probability $1 - \theta$, $z = 0$), or of the danger of consumption (with probability θ, $z = 1$). All samples realized in a play of the game are stochastically independent. Sampling is sequential and we suppose that each self may, at any moment, stop the learning process and rationally behave according to his current information. Note that by performing an infinite number of experiments, he could alternatively learn the true value of θ. For simplicity, we shall also assume that sampling is the only source of information about θ, so that in particular consumption decisions are uninformative. As already observed, this is the case if the externality from past consumption is unobservable at each date.

- *Cost of experimentation.*   We will assume that all observations are costless for the consumer. Since experimentation does not create any delay for the consumption decision, this implies that there is no cost of information acquisition for any self.

- *Information updating and perfect recall.*   After each experimentation, posterior beliefs about θ are updated according to Bayes rule. Since the consumer is viewed as the collection of all selves, it is natural to assume that each self observes the outcome of all his predecessors' sampling and therefore inherits the posterior belief of his immediate predecessor as an input for his own experimentation decision. This perfect recall feature of individual learning leads to an informational externality, which adds to the strategic conflict arising from consumption externalities.

A strategy for self-t in the experimentation/consumption game consists of two parts: first, a *sampling strategy* and second, a *consumption strategy*. Since from the perspective of any self, the only payoff relevant variable is his current belief about θ, we shall be only concerned with equilibria in Markov strategies contingent on the posterior distribution of beliefs.

- *Markov strategies.*   Basically, a Markov sampling strategy for self-t is described by a subset S_t of the space $\Pi(\pi_0)$ of posterior distributions that can be reached after a finite number of samples, given the prior distribution π_0. S_t will be interpreted as self-t's

stopping region for the stochastic process of beliefs. Moreover, since all selves face essentially the same decision problem, it is natural to focus on stationary Markov perfect equilibria (MPE) in which the sampling and consumption strategies are symmetric across all selves. It follows that all sampling is performed by self-0 on an MPE path, since there is no other information flow about θ that sampling itself and all selves stop sampling on the same region of posterior beliefs.

Admittedly, in an intra-personal game, non-Markov strategies may be used very naturally to mitigate the time inconsistency problem. However, our objective is to show that self-restriction in information gathering may *per se* be used as a self-commitment device. This point is therefore better illustrated by Markov strategies, rather than by more complex strategies.

3.3. *Complete learning versus strategic ignorance*

Our main concern is to determine whether, in this general framework, incomplete learning may occur on an MPE path. The following proposition provides a full description of the equilibria.

Proposition 1. *The experimentation/consumption game has two classes of equilibria.*

(i) *A complete learning MPE in which each self learns the true θ unless it is already known.*

(ii) *An infinite family of pure-strategy strategic ignorance MPEs in which:*

- *Incomplete learning about θ takes place with positive probability on the equilibrium path.*
- *If sampling stops after a finite number of samples, each self abstains from consuming.*
- *From any belief on, there is a strictly positive probability that the experimentation process will eventually stop.*

Proof.   See Appendix. ||

Part (i) is straightforward. Suppose that self-t anticipates that self-$t+1$ will learn the true value of θ through infinite sampling, independently of his inherited beliefs. Then, self-t exerts no informational externality on his successors, whatever his sampling strategy. As a result, he will be better off being fully informed, in order to take the current optimal consumption decision. Since the same argument can be repeated for each self, complete learning from the outset is always an MPE.

However, part (ii) states that *each self may decide to stop sampling when some posterior distribution of beliefs is reached*. The incentives to do so depend crucially on the sampling strategy followed by subsequent selves. For instance, no self wants to perform less sampling than his successors: if some information is going to be obtained by self-$t+1$, then it is in self-t's interest to acquire this information himself.[12] Now, suppose that self-t anticipates that all subsequent selves will stop the sampling process and refrain from consuming on a region $S \subset \{\pi \in \Pi(\pi_0) | E_\pi(\theta) \geq \theta^*\}$. From the above reasoning, self-t is never going to stop sampling before hitting S. If he reaches a distribution of beliefs $\pi \in S$,

12. Note that this is precisely the reason why complete learning is always an equilibrium.

he faces two options. On the one hand, he may follow the prescribed strategy, *i.e.* to stop learning (and abstain, since $E_\pi(\theta) \geqq \theta^*$). On the other hand, he may deviate and continue sampling. This deviation is profitable if it leads the agent to learn that the true θ lies below the IC region, *i.e.* in $[0, \tilde{\theta}]$, because in that case consuming at every period is optimal (see Figure 1). The deviation is not desirable if the agent learns that θ belongs to IC, since it implies that every future incarnation will inefficiently consume. Last, if the true $\theta \in [\theta^*, 1]$, it is immaterial whether the agent learns it or not (in both cases every incarnation will optimally abstain). If the precision of beliefs at π is high, the likelihood of a large shift in posteriors is low. Then, by continuing the experimentation process, self-t is much less likely to learn that θ lies below IC than to find that θ is above θ^* or, much worse, to get inefficiently stuck in the IC region.

Overall, we have the analogue of the three-period example: if, conditional on $\theta < \theta^*$ and for beliefs π, the true θ is more likely to be in the interval IC than below IC, then self-t optimally stops sampling at π, provided that no future self reinitiates the experimentation process. As the same reasoning holds for all selves, stopping when the distribution of beliefs is in the region S is an MPE.[13] Hence, despite the strategic constraint imposed on each self by his successors' experimentation, self-restriction in information gathering may still be used as a self-commitment device, thus providing a rationale for voluntary "strategic ignorance", even if no commitment to information acquisition is feasible in any period.

At this stage, we can make precise in which sense beliefs are manipulated in our model. Naturally, there cannot be an *ex ante* bias in the first-order sense, since Bayesian updating implies that the stochastic process of beliefs is a martingale. Hence, only the higher-order moments of beliefs can be manipulated. Still, in a more subtle way, our work provides strong, testable predictions for some biased beliefs phenomena. In the study by Viscusi (1990) mentioned in the introduction, the bias between subjective and objective risks assessments is systematic: the probability of a negative delayed outcome is *overestimated*. In other words, the existence of a bias is not established *ex ante*, but conditionally on the truth. Interestingly, this observed bias is in line with the predictions of our model. According to our result, if the true value of θ lies below θ^* there is a "pessimistic bias" conditionally on the truth. To see this, consider a population of time-inconsistent individuals, each of them playing some strategic ignorance MPE. At the aggregate level, a fraction of the population learn the truth and consume accordingly, while the rest of the population will remain partially uninformed, overestimate the risk, and abstain.[14]

The existence of an equilibrium where all the information is collected from the outset raises some doubts about the significance of the more appealing equilibria with incomplete learning. However, we have the following strong result.

Proposition 2. *The expected welfare of all selves is strictly higher in any strategic ignorance* MPE *than in the complete learning* MPE.

Proof. See Appendix. ‖

13. To provide the intuition, we have implicitly assumed that a deviation from the prescribed strategy leads to learning the true θ. However, the same reasoning holds for "one-shot deviations", *i.e.* deviations such that the corresponding self stops sampling if he hits S a second time.

14. Note that in our model, no interesting predictions can be drawn if the true θ lies above θ^*. overestimation or underestimation of risk is immaterial since, in equilibrium, everybody will follow the optimal behaviour, *i.e.* abstention. However, this result depends crucially on the assumption of binary consumption decisions

The intuition for this result is straightforward. Suppose indeed that self-t is strictly better off learning θ than stopping the experimentation in some region S. Then, self-t's best response to his successors stopping according to S is to become perfectly informed. However, this implies that S is not a self-sustaining MPE. In other words, the very existence of MPEs where the agent avoids collecting costless information implies that they necessary yield for any posterior beliefs and for all selves a higher expected welfare than the one obtained under full information. We can therefore justify our claim that individuals are better off acting under self-restricted information rather than under perfect knowledge.

Before concluding, several comments are in order:

• *Multiple equilibria.* The fact that no self wants to perform less sampling than his successors leads to a coordination problem in learning which results in a multiplicity of MPEs. Indeed, the complete learning equilibrium is just the limiting case in which each self anticipates that the region where his immediate successor stops sampling converges to the empty set.

• *Necessary and sufficient conditions for strategic ignorance.* When $\beta = 1$, the intrapersonal conflict of interests vanishes, so information is always valuable. More importantly, with infinite sampling opportunities and $\beta < 1$, there always exists a region where incarnations can efficiently stop the learning process and abstain. Overall, time inconsistency is a necessary and sufficient condition for strategic ignorance to be consistent with rationality.

• *Incomplete learning and consumption.* $E_\pi(\theta) \geq \theta^*$ is a necessary condition for stopping the experimentation process. Stated differently, incomplete learning is incompatible with a positive consumption in equilibrium because any self could be better off by reinitiating the sampling process and efficiently abstaining if the true θ is above θ^*. We do not emphasize this result because it heavily relies on the assumption of binary consumption decisions.

• *The variance effect.* $E_\pi(\theta) \geq \theta^*$ is not a sufficient condition for sampling to stop. Indeed, as long as the variance of beliefs remains relatively large, the risk of learning that the true value of θ falls in IC, and therefore of "being trapped" in overconsumption, is relatively low. In this case, it is optimal to pursue the experimentation even if the expected posterior is above θ^*.

• *Other sources of learning.* The results easily extend to more realistic cases in which there is exogenous learning and/or informative consumption (say, because the externality is observable). As shown in an earlier version (Carrillo and Mariotti (1998)), the trade-off implied by experimentation (optimal current consumption *vs.* learning inefficiently shared with future selves) still holds when we include other sources of information. The basic idea is the following. Each incarnation knows that, if some external information flows in, consumption may be reinitiated. However, all selves have incentives to delay this inefficient consumption as much as possible, and this is achieved through current ignorance and abstention.

4. CONCLUDING REMARKS

This paper does not pretend to be the first to highlight the strategic value of ignorance. Since Hirshleifer (1971), the idea that information may hurt has become a recurrent theme

in the game theory literature. However, all the games studied so far share some common features. First, as in Hirshleifer's example, public information may destroy mutually beneficial insurance possibilities. Second, individual incentives to gather private information may be limited by the signalling value of the actions taken by privately informed players, as in Grossman and Stiglitz (1980). By contrast, our model would be analogous to a multi person situation where the information obtained by any individual becomes automatically *public*. While this assumption is in general hard to motivate, it seems particularly natural in our intra-personal game with perfect recall. Another point of departure from related results lies in the commitment assumptions. It has been argued (*e.g.* by Crémer (1995)) that in principal–agent problems, a credible commitment of a principal not to acquire information about the agent may strengthen incentives and overwhelm the gains from better information. By contrast, a crucial feature of our strategic ignorance equilibria is that at any point in time, the agent voluntarily restricts his access to information, so that commitment to information gathering is not necessary.

We have emphasized that information may have a negative value. However, this distortion is not unidirectional: for agents with prior beliefs in IC, information is extremely valuable, since it may allow the agent to build some self-commitment power. Therefore, if information is costly, a time inconsistent individual may end up acquiring more information than his time consistent peer.

Last, our theory may be helpful to explain some puzzles in decision-making largely documented by psychologists. Indeed, our model provides a rationale for one form of cognitive dissonance, in the sense of Festinger (1957). Specifically, in the case where an irreversible investment has been already sunk, our theory predicts that a time inconsistent agent may wish to remain optimistic about the prospects attached to this investment as time comes to build on it. However, the interpretation of other forms of cognitive dissonance or self-justification (in the sense of Aronson (1972)) in terms of "rational" decision-making remains a largely open question.

APPENDIX

It is obvious from Lemma 2 that in any MPE, the consumption decision of all selves conditional on their belief π about θ is to consume if $E_\pi(\theta) < \theta^*$ and to abstain if $E_\pi(\theta) > \theta^*$. If $E_\pi(\theta) = \theta^*$, then all selves are indifferent beween consuming and abstaining, and there is no loss of generality in assuming that they abstain in this case. This yields a continuation value function

$$V(E_\pi(\theta)) = \begin{cases} 1 + \dfrac{1}{1-\delta}[\beta\delta - E_\pi(\theta)/\theta^*] & \text{if } E_\pi(\theta) < \theta^*, \\ 0 & \text{if } E_\pi(\theta) \ge \theta^*. \end{cases} \tag{A.1}$$

Thus, an MPE is characterized by a stopping region $S \subset \Pi(\pi_0)$ for the sampling process.

A *The stochastic environment*

Let $Z_x = \{0, 1\}^{\prime}$ be the observations space, and $\mathscr{F}_\sigma$ the natural σ-field on $Z_\prime$. For any $\pi \in \Pi(\pi_0)$, let μ_π be the probability measure on $[0, 1] \times Z_\prime$ induced by π and the likelihood function. For any $\pi \in \Pi(\pi_0)$, there is a limit distribution $\pi_\prime$ such that, if infinite sampling is performed from π on, then with μ_π-probability one the posterior beliefs converge to $\pi_\prime$ and $\pi_\prime = \delta_\theta$. Hence, if some self performs infinite sampling from π on, his expected continuation payoff is $\int_{[0,1] \times Z_\prime} V(E(\theta | \mathscr{F}_\sigma)(\theta, z_\sigma)) \mu_\pi(d\theta, dz_\sigma) = E_\pi(V(\theta))$. Last, identifying $S \subset \Pi(\pi_0)$ with the set of observation paths leading to it, let $\Pr_\pi(S) = \mu_\pi([0, 1] \times S)$ be the probability of hitting S starting from belief π.

B. *Proof of Proposition* 1(i).

We prove that $\varnothing$ is an MPE. Suppose that self-0 deviates from $\varnothing$ by stopping and consuming at π with $E_\pi(\theta) < \theta^*$. Since self-1 resumes sampling at π and learns the true value of θ with probability one, self-0's expected payoff from deviating is

$$U = 1 - E_\pi(\theta)/\theta^* + \pi(\theta < \theta^*)\frac{\delta}{1-\delta}[\beta - E_\pi(\theta \mid \theta < \theta^*)/\theta^*] \tag{A.2}$$

Hence $E_\pi(V(\theta)) - U = \pi(\theta \geqq \theta^*)[E_\pi(\theta \mid \theta \geqq \theta^*)/\theta^* - 1] > 0$, by the law of total expectations and the full support assumption, so that $\varnothing$ is a strict better response for self-0 at π. Suppose now that self-0 deviates from this strategy by stopping and abstaining at π with $E_\pi(\theta) \geqq \theta^*$. Since self-1 resumes sampling at π and learns the true value of θ with probability one, self-0's expected payoff from deviating is

$$U' = \pi(\theta < \theta^*)\frac{\delta}{1-\delta}[\beta - E_\pi(\theta \mid \theta < \theta^*)/\theta^*]. \tag{A.3}$$

By the law of total expectations, $E_\pi(V(\theta)) - U' = \pi(\theta < \theta^*)[1 - E_\pi(\theta \mid \theta < \theta^*)/\theta^*] > 0$, so that $\varnothing$ is a strict better response for self-0 at π. Due to the additive separability of payoffs, the same argument holds for all selves, so that $\varnothing$ is an MPE. ||

C. *Proof of Proposition* 1(ii)

We first prove that if an MPE exhibits strategic ignorance, all selves abstain from consuming if they stop learning Formally:

Lemma 3. *If $S \neq \varnothing$ is an MPE, $E_\pi(\theta) \geqq \theta^*$ for each $\pi \in S$.*

Proof. Suppose the contrary, and let $\pi \in S$ such that $E_\pi(\theta) < \theta^*$. Since all selves consume at π, the expected continuation payoff at π is $V(E_\pi(\theta))$ for any self Since $\beta < 1$, it follows that

$$E_\pi(V(\theta)) - V(E_\pi(\theta)) = \pi(\theta \geqq \theta^*)\frac{1}{1-\delta}[E_\pi(\theta \mid \theta \geqq \theta^*)/\theta^* - 1 + \delta(1-\beta)] > 0, \tag{A.4}$$

so that a strict better response for any self would be to sample infinitely from π, a contradiction. ||

For any $\pi \in \Pi(\pi_0)$ and $S \subset \Pi(\pi_0)$, let π_S be equal to the posterior distribution of beliefs at the first time S is hit by the sampling process when this happens, and to π_∞ otherwise Let $S^* = \arg\max_{S \subset \Pi(\pi_0)} E_\pi(V(E_{\pi_S}(\theta)))$ be the optimal sampling strategy from the perspective of self-0. Standard optimal stopping theory imply that S^* exists and is essentially unique. To characterize S^*, the following lemma will be helpful.

Lemma 4. *If π_0 has a continuous, full support density with respect to Lebesgue measure that is bounded away from zero, and $(\pi_n)_{n \in \mathbb{N}}$ is a sequence in $\Pi(\pi_0)$ such that $E_{\pi_n}(\theta) \downarrow \theta^*$, then $\lim_{n \to \infty} E_{\pi_n}(\theta \mid \theta < \theta^*) = \theta^*$.*

Proof. If π_0 has a continuous density f_0 with respect to Lebesgue measure, then each π_n can be identified with its density $f_n \propto \theta^n(1-\theta)^{p(n)}f_0(\theta)$, where n is the number of samples $z = 1$ and $p(n)$ is the number of samples $z = 0$. For each small $\varepsilon > 0$ and $n \in \mathbb{N}$, it is easy to see that

$$E_{\pi_n}(\theta \mid \theta < \theta^*) > (\theta^* - \varepsilon)\left(1 + \frac{\int_0^{\theta^*-\varepsilon} \theta^n(1-\theta)^{p(n)}f_0(\theta)d\theta}{\int_{\theta^*-\varepsilon}^{\theta^*} \theta^n(1-\theta)^{p(n)}f_0(\theta)d\theta}\right)^{-1} \equiv I(n, \varepsilon). \tag{A.5}$$

We show that $\lim_{n \to \infty} I(n, \varepsilon) = \theta^* - \varepsilon$. Since f_0 is continuous and bounded away from zero,

$$\frac{\int_0^{\theta^*-\varepsilon} \theta^n(1-\theta)^{p(n)}f_0(\theta)d\theta}{\int_{\theta^*-\varepsilon}^{\theta^*} \theta^n(1-\theta)^{p(n)}f_0(\theta)d\theta} \leq \frac{\max_{\theta \in [0, \theta^*-\varepsilon]}f_0(\theta)}{\min_{\theta \in [\theta^*-\varepsilon, \theta^*]}f_0(\theta)}\frac{\int_0^{\theta^*-\varepsilon} \theta^n(1-\theta)^{p(n)}d\theta}{\int_{\theta^*-\varepsilon}^{\theta^*} \theta^n(1-\theta)^{p(n)}d\theta}. \tag{A.6}$$

Now assume that $\lim_{n \to \infty} n/(n + p(n)) = \theta^*$. Since the mapping $\theta \mapsto \theta^n(1-\theta)^{p(n)}$ is increasing over $[0, n/(n+p(n))]$, it follows that for n large enough, $\int_0^{\theta^*-\varepsilon} \theta^n(1-\theta)^{p(n)}d\theta < (\theta^*-\varepsilon)^{n+1}(1-\theta^*+\varepsilon)^{p(n)}$. From this and Jensen's inequality

$$\frac{\int_0^{\theta^*-\varepsilon} \theta^n(1-\theta)^{p(n)}d\theta}{\int_{\theta^*-\varepsilon}^{\theta^*} \theta^n(1-\theta)^{p(n)}d\theta} < \frac{\theta^*-\varepsilon}{\varepsilon^2}\int_{\theta^*-\varepsilon}^{\theta^*}\left(\frac{\theta^*-\varepsilon}{\theta}\right)^n\left(\frac{1-\theta^*+\varepsilon}{1-\theta}\right)^{p(n)}d\theta. \tag{A.7}$$

We now show that the r.h.s. in (A.7) goes to zero as n goes to infinity. For each $\theta \in (\theta^* - \varepsilon, \theta^*]$, we have $\lim_{n \to \infty} \ln(\theta^* - \varepsilon) - (p(n)/n) \ln(1 - \theta^* + \varepsilon) - \ln(\theta) - (p(n)/n) \ln(1 - \theta) = \ln(\theta^* - \varepsilon) + (1 - \theta^*)/\theta^* \ln(1 - \theta^* + \varepsilon) - \ln(\theta) - (1 - \theta^*)/\theta^* \ln(1 - \theta) \equiv l(\theta, \varepsilon)$. Moreover, the mapping $\theta \mapsto \ln(\theta) + (1 - \theta^*)/\theta^* \ln(1 - \theta)$ is increasing on $(0, \theta^*]$. Hence $l(\theta, \varepsilon) < 0$ for each $\theta \in (\theta^* - \varepsilon, \theta^*]$. Let $\theta \in (\theta^* - \varepsilon, \theta^*]$, and $\eta > 0$ such that $l(\theta, \varepsilon) + \eta < 0$. There exists $n(\theta, \eta)$ such that for all $n \geq n(\theta, \eta)$, $((\theta^* - \varepsilon)/\theta)^n ((1 - \theta^* + \varepsilon)/1 - \theta)^{p(n)} < e^{n(l(\theta, \varepsilon) + \eta)}$ and the result follows. For each $n \geq 0$, the mapping $\theta \mapsto \theta^n (1 - \theta)^{p(n)}$ is increasing on $[0, n/(n + p(n))]$, and decreasing on $[n/(n + p(n)), 1]$. Hence two cases may arise. Either $n/(n + p(n)) \in [\theta^*, 1]$, in which case the integrand is bounded by one. Or $n/(n + p(n)) \in (\theta^* - \varepsilon, \theta^*)$, however, if n is large enough, $(\theta^* - \varepsilon)^n (1 - \theta^* + \varepsilon)^{p(n)} < (\theta^*)^n (1 - \theta^*)^{p(n)}$, and the integrand is again bounded by one. It follows from Lebesgue's dominated convergence theorem that the integral on the r h s of (A 7) converges to zero as n goes to infinity. From (A.6), we have $\lim_{n \to \infty} I(n, \varepsilon) = \theta^* - \varepsilon$. From (A.5), it follows that for each $\varepsilon > 0$, there exists $n(\varepsilon)$ such that for each $n \geq n(\varepsilon)$, $\theta^* \geq E_{\pi_n}(\theta | \theta < \theta^*) > I(n, \varepsilon/2) > \theta^* - \varepsilon/2 - \varepsilon/2$. Hence, to complete the proof, we must only show that if $\lim_{n \to \infty} E_{\pi_n}(\theta) = \theta^*$, then $\lim_{n \to \infty} n/(n + p(n)) = \theta^*$. Using exactly the same techniques as above, one can show that if $\lim_{n \to \infty} n/(n + p(n)) = \bar{\theta} \neq \theta^*$, then $\lim_{n \to \infty} E_{\pi_n}(\theta) = \bar{\theta}$, a contradiction. ||

An an immediate consequence of Lemma 4, we obtain

Lemma 5.   *If π_0 has a continuous, full support density with respect to Lebesgue measure that is bounded away from zero, then $S^* \neq \varnothing$ and $\mathrm{Pr}_\pi(S^*) > 0$ for each $\pi \in \Pi(\pi_0)$*

Proof.   As π_0 has a continuous, full support density on $[0, 1]$, there exists a sequence $(\pi_n)_{n \in \mathbb{N}}$ such that $E_{\pi_n}(\theta) \downarrow \theta^*$. Since $\beta > 1$, it follows from Lemma 4 that there exists $n_0 \in \mathbb{N}$ such that $E_{\pi_n}(\theta) > \theta^*$ and $[1 + \delta(\beta - 1)]\theta^* < E_{\pi_n}(\theta | \theta < \theta^*) < \theta^*$ for each $n \geq n_0$. Now, suppose that $S^* = \varnothing$. Then, for any $\pi \in \Pi(\pi_0)$ such that $E_\pi(\theta) \geq \theta^*$, one must have $E_\pi(V(\pi_\varnothing(\theta))) = E_\pi(V(\theta)) \geq 0$. However, since for $n \geq n_0$, $[1 + \delta(\beta - 1)]\theta^* < E_{\pi_n}(\theta | \theta < \theta^*)$, one has $E_{\pi_n}(V(\theta)) < V(E_{\pi_n}(\theta)) = 0$, a contradiction. It follows in particular that for all $n \geq n_0$, there is a belief $\pi_n^* \in S^*$ that can be reached in a finite number of sample from π_n. Hence $\mathrm{Pr}_\pi(S^*) > 0$ for each $\pi \in \Pi(\pi_0)$, as claimed. ||

By Lemma 3, it is easy to see that for each $\pi \in S^*$, $E_\pi(\theta) \geq \theta^*$. The existence of a strategic ignorance MPE follows then from the next lemma.

Lemma 6. S^* *is an* MPE.

Proof.   Suppose that all selves-t, $t \geq 1$ follow the sampling strategy S^*. Then, since by construction S^* is optimal from the perspective of self-0, a best response for self-0 is to play S^* as well. Since S^* maximizes the continuation payoff of any self, it follows that S^* is an MPE. ||

To complete the proof, we must show that there exists an infinite family of strategic ignorance MPE. Note first that for any $\pi \in S^*$, $\{\pi\}$ is a strategic ignorance MPE, since each self prefers to stop sampling at π and consistently abstain rather than to learn the true value of θ by infinite sampling. Note however that the stopping region $\{\pi\}$ cannot be reached with positive probability from all posterior beliefs in $\Pi(\pi_0)$ However, using the same techniques as above, it is easily checked that by removing from S^* all the posterior distributions of beliefs that can be reached in less than a finite and given number of samples, one obtains again a strategic ignorance MPE ||

D. *Proof of Proposition 2*

Let $S \neq \varnothing$ be an MPE. It is obvious that for any $\pi \in S$, the expected payoff for any self at π, *i.e.* 0, must be at least as large as the complete learning continuation payoff $E_\pi(V(\theta))$, otherwise some self would have an incentive to deviate and learn the true value of θ. By the law of total expectations, it follows that at any $\pi \in \Pi(\pi_0)$, the expected payoff to any self from playing S is at least as large as his expected payoff from playing $\varnothing$. Next, for any MPE $S \neq \varnothing$, one can show in exactly the same way as in the proof of Lemma 5 that there exists a $\pi \in S$ such that $E_\pi(\theta | \theta < \theta^*) > [1 + \delta(\beta - 1)]\theta^*$, so that $E_\pi(V(\theta)) < V(E_\pi(\theta)) = 0$ and the result follows ||

REFERENCES

AGHION, P., BOLTON, P., HARRIS, C. and JULLIEN, B. (1991), "Optimal Learning by Experimentation", *Review of Economic Studies*, **58**, 621–654.

AINSLIE, G. (1975), "Specious Reward: a Behavioral Theory of Impulsiveness and Impulse Control", *Psychological Bulletin*, **82**, 463–509.

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

AKERLOF, G. A. (1991), "Procrastination and Obedience", *American Economic Review*, **81**, 1–19.

ARONSON, E. (1972) *The Social Animal* (San Francisco: W. H. Freeman).

BECKER, G. S. and MURPHY, M. K. (1988), "A Theory of Rational Addiction", *Journal of Political Economy*, **96**, 675–700.

CAILLAUD, B., COHEN, D. and JULLIEN, B. (1996), "Toward a Theory of Self-Restraint" (Manuscript, CEPREMAP, Paris).

CARRILLO, J. D (1998), "Self-Control, Moderate Consumption, and Craving" (CEPR D.P. 2017).

CARRILLO, J. D. and MARIOTTI, T. (1998), "Strategic Ignorance as a Self-Disciplining Device", (Manuscript, ECARE, Brussels).

CRÉMER, J. (1995), "Arm's Length Relationships", *Quarterly Journal of Economics*, **110**, 275–295.

EASLEY, D. and KIEFER, N. M (1988), "Controlling a Stochastic Process with Unknown Parameters", *Econometrica*, **56**, 1045–1064.

FESTINGER, L. A. (1957) *A Theory of Cognitive Dissonance* (Stanford: Stanford University Press).

GROSSMAN, S. J. and STIGLITZ, J. (1980), "On the Impossibility of Informationally Efficient Markets", *American Economic Review*, **70**, 393–408.

HIRSHLEIFER, J. (1971), "The Private and Social Value of Information and the Reward from Incentive Activity", *American Economic Review*, **61**, 561–574.

KELLER, G. and RADY, S. (1999), "Optimal Experimentation in a Changing Environment", *Review of Economic Studies*, **66**, 475–507.

LAIBSON, D. I. (1996), "Hyberbolic Discount Functions, Undersaving, and Savings Policy" (Manuscript, Harvard University).

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

LOEWENSTEIN, G and PRELEC, D. (1992), "Anomalies in Intertemporal Choice. Evidence and an Interpretation", *Quarterly Journal of Economics*, **107**, 573–598.

McLENNAN, A. (1984), "Price Dispersion and Incomplete Learning in the Long Run", *Journal of Economic Dynamics and Control*, 7, 331–347.

MAZUR, J. E. (1987), "An Adjustment Procedure for Studying Delayed Reinforcement", in Commons, M. L., Mazur, J. E., Nevins, J. A. and Rachlin, H. (eds.), *Quantitative Analysis of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value* (Hillsdale: Ballinger).

O'DONOGHUE, T. and RABIN, M. (1998), "Addiction and Self Control" (Manuscript, University of California, Berkeley).

O'DONOGHUE, T. and RABIN, M. (1999), "Doing it Now or Later", *American Economic Review*, **89**, 103–124.

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

ROTHSCHILD, M. (1974), "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory*, 9, 185–202.

STROTZ, R. H. (1956), "Myopia and Inconsistency in Dynamic Utility Maximisation", *Review of Economic Studies*, **23**, 166–180.

THALER, R. H. (1981), "Some Empirical Evidence on Dynamic Inconsistency", *Economic Letters*, **8**, 201–207.

THALER, R. H. (1990), "Saving, Fungibility, and Mental Accounts", *Journal of Economic Perspectives*, **4**, 193–205.

THALER, R. H. and SHEFRIN, N. M. (1981), "An Economic Theory of Self-Control", *Journal of Political Economy*, **89**, 392–406.

VISCUSI, W. K. (1990), "Do Smokers Underestimate Risks?", *Journal of Political Economy*, **98**, 1253–1269.