The Evolution of Self-Control in the Brain

David Jimenez-Gomez*

January 10, 2023

Abstract

I present a theoretical framework that shows how limited self-control could have evolved as a mechanism to make humans behave against their own self-interest. I analyze the evolution of self-control in a principal-agent framework with two agents, System 1 and System 2, that represent the automatic and cognitive processes within the human mind, respectively. Based on the relevant evidence, I assume that System 2 has access to private information, but its utility cannot depend on all the relevant information. The principal can achieve the asymptotically optimal outcome by biasing the utility of System 2 (from which an endogenous conflict emerges) and simultaneously endowing it with a limited amount of self-control. The model explains several empirical properties of self-control (observed in experiments), and sheds light on its welfare implications.

JEL: D60, D90, C72, D81.

Keywords: self-control, neuroeconomic theory, evolution of preferences.

---

*University of Alicante. Email: davidjimenezgomez@ua.es. I thank Daron Acemoglu, Benjamin Brooks, Luigi Butera, Dana Chandler, Carlos Cueva, Alexander Franklin, Max Kleiman-Weiner, Miguel Risco Bermejo, Adam Sanjurjo, Hugo Sonnenschein, Josh Tenenbaum, Citlalli Vivar, Nils Wernerfelt, the participants of seminars at Josh Tenenbaum's group at the MIT Brain and Cognitive Sciences Department, and at the University of Murcia, as well as anonymous referees, for their comments. I am grateful for the funding from the Ministry of Science and Innovation and the European Regional Development Fund (FEDER), through the PID2019-107081GB-I00 project; from the Ministry of Science and Innovation, through the project "Nudging applied to the improvement of regulation" (RED2018-102761-T); from the Valencian Community, through the Prometeo program PROMETEO/2021/073.

1 Introduction

Self-control is the ability to control or override one's thoughts, emotions, urges, and behavior (Gailliot *et al.*, 2007). Self-control is a better predictor of college grades than IQ or SAT score (Duckworth and Seligman, 2005), higher levels of self-control are associated with lower incidence of mental illness and alcohol abuse (Tangney *et al.*, 2004), and people who have more self-control are better educated, wealthier and healthier (Heatherton and Wagner, 2011; Mischel, 2014). Self-control seems therefore like a primary target for policy interventions, because increasing self-control can increase the health, education, and productivity of individuals (Blattman *et al.*, 2017; Alan *et al.*, 2019). Crucially, if self-control is so beneficial, and given that humans are the result of an evolutionary process, this begs the question: why have we not evolved a perfect self-control (Hayden, 2019)? It would seem a priori that having more self-control would lead to more successful individuals, who would be better able to survive and reproduce (Baumeister and Tierney, 2011; McGonigal, 2011). In this paper I argue that there is actually an optimal amount of self-control from the point of view of genetic fitness: in other words, too much self-control is detrimental for the survival and reproduction of our genes, and that is why we have evolved a limited self-control.[1]

I consider a framework where an individual makes decisions that are generated from the interaction between two brain systems: System 1 (automatic or "hard wired") and System 2 (conscious and reflexive, Kahneman, 2011; Cerigioni, 2021). The individual has a series of K opportunities to either cheat or to respect a social norm.[2] The human genes are subject to natural selection, that maximizes their replication to the next generation (i.e. quality-adjusted

[1] The notion that impulsivity could be adaptive has been proposed by several authors (Sozou, 1998; Stephens and Anderson, 2001; Stephens, 2002; Kacelnik, 2003; Fawcett *et al.*, 2012). However, the present model proposes a novel mechanism, in that it is not impulsivity itself that is adaptive, but having limited amount of self-control that forces the individual to behave impulsively in the right circumstances (i.e. when the fitness gains would be higher).

[2] Cheating in the presence of social norms has been widely studied, for example in the context of academic cheating, Teixeira and Rocha, 2010; scientific fraud, Necker, 2016; and tax evasion, Slemrod, 2007. Social norms have been extensively studied (Schultz *et al.*, 2007; Postlewaite, 2011), as well as their economic applications (Allcott, 2011). Cheating in social contexts has also been studied in evolutionary psychology and cognitive science Cosmides (1989); Gigerenzer and Hug (1992); Cosmides and Tooby (1992) and behavioral economics (Shu *et al.*, 2012).

number of offspring, what is known as *genetic fitness*), and each of the two actions (respecting norms vs. cheating) has fitness gains that are stochastic. I formalize this setup in a principal-agent model, where the human genes are represented by the Principal, and the human individual is represented by two different players: System 1 and System 2 (the "multiple-selves" approach to self-control follows a long tradition; Schelling 1978; Thaler and Shefrin 1981; Bernheim and Rangel 2004; Benhabib and Bisin 2005; Fudenberg and Levine 2006; Brocas and Carrillo 2008). Following Kahneman (2011), I consider that System 1 is genetically hard-wired by the Principal, and will be in charge of implementing actions "automatically"; while System 2 can be identified with higher cognitive processes.[3] The Principal chooses the utility function of Systems 1 and 2 (subject to constraints) i.e. utilities are endogenous and dependent on the genes, an assumption that is standard in the literature on the *evolution of preferences* (Robson, 2001; Samuelson, 2004; Samuelson and Swinkels, 2006; Rayo and Becker, 2007; Netzer, 2009; Robson and Samuelson, 2011; Rayo and Robson, 2016).

If there were no further constraints, the Principal could achieve the fitness-maximizing outcome simply by choosing an unbiased utility for System 2 (Robson and Samuelson, 2010), in which case self-control would play no role, as there would be no conflict of interest between the Principal and System 2. However, based on the evidence from neuroscience that I review in Section 2.3, I assume that the Principal cannot include all the relevant variables as arguments in System 2's utility function (a similar assumption is made by Brocas and Carrillo, 2008; Rayo and Robson, 2016). This means that System 2 has superior information (relative to both Principal and System 1) with respect to idiosyncratic details of the social norms that operate in the individual's environment, and therefore the Principal has an incentive to extract this information from System 2. The Principal can do this by linking together the outcomes that result from the K decisions taken by System 2, generating accountability across decisions (Jackson and Sonnen-

[3]In particular, System 2 can be interpreted as the brain area known as the dorsolateral prefrontal cortex (dlPFC), that is responsible for planning (Elliott, 2003) and self-control (Knoch *et al.*, 2006; Hare *et al.*, 2009; Figner *et al.*, 2010). I provide the relevant neuroscientific evidence in Section 2.3. However, despite being identifiable with certain brain areas, I would like to follow Kahneman (2011) in emphasizing that Systems 1 & 2 are not necessarily two discrete systems, but should be understood as metaphors for complex organizations within the brain.

schein, 2007). In particular, System 2 can send messages to System 1, and each message should be interpreted as a possibility for System 2 of exerting self-control (the larger the message, the more self-control is exerted). By imposing a "self-control budget", the Principal ensures that System 2 exerts self-control proportionally to how valuable it is to follow the social norm in each specific instance, in such a way that System 1 will choose cheating when the fitness gains from cheating exceed the gains from following the social norm for that particular instance. Therefore, I argue that *limited self-control evolved as a mechanism to make humans behave against their own self-interest in order to increase genetic fitness.*[4]

The main result of the paper is that the Principal can achieve the (asymptotically) fitness-maximizing outcome by endowing System 2 with a biased utility, while simultaneously restricting its autonomy via a "self-control budget" (Proposition 1). In other words, natural selection produces a utility function that is biased in a way that gives the human individual enough autonomy to adapt to a changing environment, while simultaneously restricting the individual's autonomy to make her comply with genetic fitness (rather than utility) maximization.[5] This provides an explanation for the apparent paradox that human's lack of perfect self-control is detrimental for the individual, and yet resulted from an evolutionary process that maximized genetic fitness.[6]

The model presented in this paper can also explain several other interesting facts about self-control. A straightforward implication is that when someone uses self-control for a task, she will have less amount of willpower left for subsequent tasks (Proposition 2), a fact observed

[4]Understanding the individual's self-interest as the evolved utility function of System 2, or possibly a convex combination of the utilities of Systems 1 and 2. I discuss this in further detail in Section 4.

[5]By a biased utility, what I mean is that the utility of System 2 does not correspond with (expected) genetic fitness. While the Principal could always choose an unbiased utility function for System 2 (in which case no conflict would exist), that would not be fitness-maximizing, because the Principal would not be able to extract all the relevant information from System 2. Similar ideas, in which the Principal selects a biased Agent, can be found in the literature: in Aghion and Tirole (1997) the Principal gives control to a biased Agent to generate incentives to collect information, and in Che and Kartik (2009) a Principal who can choose from a pool of Agents (including unbiased ones), will choose a biased Agent while retaining control so that the Agent has an incentive to collect information in order to convince the Principal.

[6]Brocas and Carrillo (2008) proposed a model of the brain in which there is asymmetric information between different brain systems. Using this framework, they obtain a result (Proposition 4) in which the Principal sets a cap on the amount of the "tempting good" that System 2 can consume. Despite the similarity between the results, they assume a benevolent Principal (who maximizes welfare) and faces self-control internalities as an external constraint, whereas I am interested precisely in how those internalities arose in the first place through natural selection.

in experiments (Baumeister and Tierney, 2011). Another observed phenomenon is that when a person exercises self-control regularly her "self-control budget" grows over time (Oaten and Cheng, 2006b, 2007). This is because if the environment is changing over time in a structured manner, the Principal can set a budget that adapts over time to better match the environment (Proposition 3). Finally, it has been observed that when people have a low level of glucose in the blood, they exert less self-control (Gailliot *et al.*, 2007; Kurzban, 2010). I show that this can be true even if glucose is only monitored, not consumed (Proposition 4). The model also has welfare implications. I show that the self-control budget is smaller than the socially optimal one (Proposition 5). Moreover, many relevant behaviors today are associated with self-control: smoking, alcohol consumption, failing to exercise, not saving enough, etc. (Baumeister *et al.*, 1994). However, these are modern behaviors that either did not exist or were not relevant during human evolutionary history (what is known as *evolutionary mismatch*). All of these behaviors are influenced by poor self-control, and I show that because the current environment is more tempting than the environment in which our genes evolved, people will not have enough self-control to make the correct decisions (according to their own utility function), what entails a reduction in welfare (Proposition 6).

This paper is connected to several literatures. Firstly, it belongs to the literature on the evolution of preferences (Robson, 2001; Samuelson, 2004; Samuelson and Swinkels, 2006; Rayo and Becker, 2007; Netzer, 2009; Robson and Samuelson, 2011; Rayo and Robson, 2016). This literature analyzes the evolution of certain traits in a Principal-Agent model, in which the Principal represents the human genes subject to natural selection, and has been quite successful at providing explanations for a variety of features in human behavior: why people have preferences instead of automatic behaviors (Robson, 2001), conspicuous consumption (Samuelson, 2004), utility dependence on unchosen alternatives (Samuelson and Swinkels, 2006), hedonic adaptation (Rayo and Becker, 2007), time and risk preferences (Netzer, 2009), and hedonic forecast bias (Robson and Samuelson, 2011); see Robson and Samuelson (2010) for a review and an in-depth justification of the framework. Particular attention deserve articles that have analyzed the evolution of intertemporal preferences in humans (Sozou, 1998; Dasgupta and Maskin, 2005; Robson and Samuelson, 2007); the main difference between these articles and the present model is that they have no conflict of interest between principal and agent (and hence intertemporal preferences simply arise as the optimal solution of a given statistical problem), whereas in the present model the (endogenous) principal-agent conflict is at the core of the explanation for self-control. This paper is also connected to the literature on decision linkage, in which a principal can discipline an agent by linking together several of the agent's choices: while the principal cannot observe particular realizations of some important variable, she knows its distribution, and can therefore incentivize the agent by linking together several of the agent's choices (Jackson and Sonnenschein, 2007; Frankel, 2014). For example, in a school in which each teacher has an incentive to inflate grades, the school principal can still extract the teacher's knowledge about the students by imposing a cap on the number of "A"s the teacher can give (Frankel, 2014), and I propose a similar mechanism in which the Principal endows System 2 with a self-control budget to link the decisions and discipline its choices.[7] This paper also belongs to a literature that consider dual decision processes (Cerigioni, 2021), in particular when such processes are applied to study self-control (Schelling 1978; Thaler and Shefrin 1981; Bernheim and Rangel 2004; Benhabib and Bisin 2005; Fudenberg and Levine 2006; Brocas and Carrillo 2008, 2021); the dual Systems model has also a long tradition in cognitive science and psychology (Sanfey and Chang, 2008; Kahneman, 2011).

2 The model

This section presents the baseline model (proofs of all results can be found in the Appendix). There are three players: the Principal, who stands for the human genes (that are subject to natural selection), System 1 (that can be seen as a genetically-programmed automatic and un-

[7]This paper is also related to a literature on "veto-based delegation", in which the principal chooses a default option and allows the agent to choose the action only within a certain set of options; otherwise the default decision is implemented (Mylovanov, 2008; Alonso and Matouschek, 2008). In the present model, it is System 1 who takes the "default action" in lieu of the Principal.

conscious module or "self" in the human brain, that executes the Principal's planned actions), and System 2 (that represents a module or "self" that is responsible for conscious decisions in the human brain).[8] The Principal's objective is to maximize genetic fitness y, interpreted as quality-adjusted offspring. There are K decisions to be made, and these should be understood as cheating opportunities (that provide a private benefit at the expense of breaking a social norm), and I will consider the case when K grows large. At each of those decision points, there are two possible actions available, $a \in \{1, 2\}$. Action 1 represents cheating: an action that has the potential to yield short-term fitness gains at the expense of the future (due to punishment from the individual's social group, Buss, 2015, Ch. 6).[9] Action 2, on the contrary, is to be interpreted as respecting social norms. Genetic fitness y is given by y^1 when action $a = 1$ chosen, and y^2 when action $a = 2$ is chosen. Both y^1 and y^2 are random and independent, as described below.

Variable y^1 is generated randomly according to distribution F, which I assume to be strictly concave and differentiable and with support in $\mathbb{R}_+$ (an exponential distribution, for example, satisfies this property, which makes the first-order conditions in System 2's problem sufficient for an optimal solution). The Principal and System 1 (but not System 2) observe y^1, which should be interpreted as the benefit from transgressing the social norm, but that is difficult for the individual to understand (for example, the value of mating at that particular moment, that can depend on the level of sexual hormones in the body, the menstrual cycle, etc).[10] System 2 (but not the Principal or System 1) observes y^2, which should be interpreted as the importance of respecting the social norms specific to the complex social environment of the individual.[11]

[8]Even though I consider the human genes (subject to the evolutionary process) as the Principal in this model, no "design" or teleology is implied. See Robson and Samuelson (2010) for a review paper on the literature that takes the approach of modeling the evolutionary process in a Principle-Agent relationship. While not a perfect match, System 2 could be seen as the "planner", and System 1 as the "doer", of planner-doer models (Thaler and Shefrin, 1981; Fudenberg and Levine, 2006).

[9]The main framework presented in this paper is static, and therefore the intuition of "short-term" gains and "long-term" costs is provided only for ease of interpreting the model. See Section 5 for a dynamic version of the main framework.

[10]This characteristic should be understood as "bottom up", i.e. outside of the awareness of the individual, but still able to be encoded by regions such as the orbitofrontal cortex in primates (OFC, Watson and Platt, 2012), and later passed on to the vmPFC and the comparator regions that I identify with System 1 (see Section 2.3 for a discussion of the relevant evidence from neuroscience).

[11] Social norms are difficult to encode in genes, because they change too fast with respect to the natural environment (although culture has undoubtedly affected the evolutionary process Boyd and Richerson, 1985).

Variable y^2 is distributed according to distribution G with support in $\mathbb{R}_+$. Moreover, y^1 and y^2 are independent. Therefore, the fitness maximizing action is to follow the social norm ($a = 2$) if and only if $y^2 \geq y^1$.

The Principal chooses the utility function $u(\cdot)$ of System 2, subject to constraints that I describe below. This follows the previous literature on the evolution of preferences, where the Principal maximizes fitness by choosing System 2's utility function (Robson and Samuelson, 2010).[12] I depart from this literature in two important aspects (for which I provide evidence from neuroscience in Section 2.3 below). The first is that, although the Principal observes y^1, it cannot make System 2's utility contingent on it (this is similar to the assumption of asymmetric information in the brain by Brocas and Carrillo, 2008).

Assumption 1. *The Principal, and Systems 1 & 2 have the following restrictions:*

1. *The Principal and System 1 observe y^1, but not y^2.*

2. *System 2 observes y^2, but not y^1.*

3. *The utility function of System 2, $u(a, y^2)$, cannot depend on y^1.*

Another departure from the previous literature on the evolution of preferences, stems from the way in which the actions are implemented. As I mentioned above, there are K decisions that need to be taken. In the present model, the incentives for System 2 of the K different choices are linked together (with $K > 1$, as opposed to $K = 1$ in the previous literature). System 2, after observing (independent and identically distributed) realizations $(y_1^2, \ldots, y_K^2)$, sends messages $(m_1, \ldots, m_K) \in \mathbb{R}_+^K$ to System 1 (each message must be non-negative, interpreted as the level of

Therefore, natural selection can design a utility function that yields payoffs for following the social norms, even if the specific social norms are only observed by the human individual.

[12] Robson (2001) showed that the reason a utility function would evolve in the first place is because it allows for flexibility in a fast-changing environment. Since social competition was one of the main causes of the evolution of intelligence, and the social environment can be very fast-changing, it is therefore not surprising that I take the choice of utility function by the Principal as the starting point of my analysis. Each paper in the literature studies a particular phenomenon by selecting which constraints the Principal faces. For example, Rayo and Becker (2007) derive the fact that happiness adapts over time (the phenomenon known as hedonic adaptation) from two constraints: utility is bounded, and System 2 has a threshold of perception under which cannot differentiate two alternatives with similar utility.

self-control exterted). Upon observing each message m_i, System 1 forms beliefs about the value of y_i^2, and then implements the fitness-maximizing action a_i, for $1 \leq i \leq K$.[13] Let $\beta(m_i)$ denote the expectation of y_i^2 after observing m_i, according to System 1's beliefs (I will refer to $\beta(m_i)$ as System 1's beliefs from now on).

Assumption 2. *System 2 sends messages $m_1, \ldots, m_K$ over K different actions to System 1, who then chooses actions $a_1, \ldots, a_K$, in such a way as to maximize fitness according to its beliefs.*

Assumption 3. *System 1's beliefs are exogenous (the Principal takes them as given), and $\beta(m)$, is increasing, concave, and differentiable.*[14]

Note that Assumption 2 implies that action $a_i = 2$ will be chosen if and only if $\beta(m_i) > y_i^1$, as System 1 maximizes fitness (given its beliefs). Finally, I assume that the Principal has an extra way of disciplining System 2: the Principal can impose a "budget" B, such that System 2 can only send messages within the budget.

Assumption 4. *The messages $m_1, \ldots, m_K$ must be such that $\frac{1}{K}\sum_{i=1}^{K} m_i \leq B$, where B is chosen by the Principal.*

In summary, the timing of decisions in the model is as follows.

1. The Principal chooses the utility function $u(a, y^2)$ for System 2, and budget B.
2. Random variables $(y_1^1, \ldots, y_K^1)$ and $(y_1^2, \ldots, y_K^2)$ are realized (i.i.d.), according to distributions F and G respectively.
3. System 2 observes the y^2 realizations and sends messages $(m_1, \ldots, m_K)$ to System 1.
4. System 1 observes the y^1 realizations and System 2's messages, and chooses actions $(a_1, \ldots, a_K)$.

[13] Previous models in the "evolution of preferences" literature assume that although the Principal can choose System 2's utility function, it is System 2 who ultimately has all the control on which action becomes implemented. Instead, I assume that the Principal retains de facto "veto power" (Mylovanov, 2008), through the intervention of System 1, and only allows System 2 to implement her preferred actions under certain conditions, which will become clear below.

[14] As it will become clear in Proposition 1 below, the Principal can obtain the first best asymptotically for any beliefs $\beta(m)$ that satisfy Assumption 3. Therefore, the outcome would be identical if I assumed that the Principal could choose $\beta(m)$ directly.

5. Fitness and utility are then obtained by Principal, System 1 and System 2.

Example 1. *To illustrate the functioning of the model, let's consider an example with* $K = 2$*, i.e. when only two actions need to be taken. Since System 2 only observes* y_1^2 *and* y_2^2*, then the Principal can extract information from System 2 by making its utility* $u(1, y^2) = 0$ *and* $u(2, y^2) = y^2$*, and setting* $B = 1/2$*, so that* $\frac{1}{2}(m_1 + m_2) \leq \frac{1}{2}$*. Assume that* y^1 *is distributed according to a geometric distribution with parameter* $\lambda = 1$*, and therefore the CDF is* $F(y^1) = 1 - e^{-y^1}$*. Suppose for simplicity that System 1 takes messages at face value, so that* $\beta(m_k) = m_k$*. Since System 1 will choose* $a_k = 2$ *if and only if* $\beta(m_k) = m_k \geq y_k^1$*, System 2 maximizes* $F(m_1)y_1^2 + F(m_2)y_2^2$*, or equivalently:*

$$\max_{m_1} [1 - e^{-m_1}]y_1^2 + [1 - e^{m_1 - 1}]y_2^2,$$

the solution of which is $m_1 = \frac{\log(y_1^2/y_2^2)+1}{2}$*. Note that System 1 can learn the relative benefit of choosing action 2 for* $k = 1$ *vs.* $k = 2$*, i.e. the ratio* y_1^2/y_2^2*, but is unable to infer the absolute values. The rest of the paper shows the Principal's optimal choice for System 2's utility and self-control budget, such that the maximal information from System 2 can be extracted.*

Assumptions 1-4 make this model depart from others in the literature of the evolution of preferences. I define the following solution concept of asymptotic equilibrium, that considers the players' behavior as K grows large, inspired by a similar solution concept in Jackson and Sonnenschein (2007). The reason to take K large is that the incentives generated by tying the decisions together grow stronger with the number of decisions K. Before defining the main concept of an asymptotic equilibrium, however, we need to define the concept of a K-equilibrium.

Definition 1. *Given a utility function* $u(a, y^2)$ *for System 2, beliefs* β *for System 1, and a budget* B*, a* K***-equilibrium*** *is given by actions* $a_1, \ldots, a_K$*, and messages* $m_1, \ldots, m_K$*, such that:*

- ◇ *System 1 chooses actions* $a_1, \ldots, a_K$*, such that it maximizes fitness* $\sum_{i=1}^{K} y_i^{a_i}$*, given its beliefs.*

◇ *System 2 choses* $m_i, \ldots, m_K$ *such that it maximizes* $\sum_{i=1}^{K} \mathbb{E}[u(a_i, y_i^2)]$*, subject to System 1's behavior and to* $\frac{1}{K}\sum_{i=1}^{K} m_i \leq B$*.*

Definition 2. *An **asymptotic equilibrium** is given by a utility function* $u(a, y^2)$ *and a budget* B *such that the Principal obtains optimal fitness asymptotically, i.e.*

$$\lim_{K\to\infty} \mathbb{P}\left[(a_i)_{i=1}^K = \arg\max \sum_{i=1}^{K} y^{a_i}\right] = 1,$$

subject to the actions of Systems 1 and 2 constituting a K*-equilibrium for* $u(a, y^2)$ *and* B*, for all* K*.*

In other words, an asymptotic equilibrium is similar to a Bayesian Nash Equilibrium, except that the requirement on System 1's beliefs, and on the Principal's optimality, are asymptotic. The asymptotic equilibrium should be understood as an approximation for a finite, but large enough, number of decisions taken by the individual. This idea is similar to the way the normal distribution is often used as an approximation, by the central limit theorem, when the number of observations ir large enough (i.e. more than thirty). Thus, the asymptotic equilibrium should be taken as an approximation, up to a small and vanishing error, of the optimal behavior of the Principal (i.e. natural selection), even for large but finite K.

2.1 System 2's problem

Recall that System 2 observes $y_1^2, \ldots, y_K^2$ and then chooses messages $m_1, \ldots, m_K$, such that $m_i \geq 0$ and $\frac{1}{K}\sum_{i=1}^{K} m_i \leq B$. From the point of view of System 2, whether System 1 chooses $a_i = 2$ conditional on sending message m_i is a random event, that has probability: $\mathbb{P}[a_i = 2] = \mathbb{P}[\beta(m_i) > y_i^1] = F(\beta(m_i))$. Because the Principal chooses a utility function $u(a, y^2)$ for System 2, the Principal can simply normalize the utility $u(1, y^2) = 0$, and then consider a function $v(y^2) = u(2, y^2)$. This implies that $(1 - F(\beta(m_i)))u(1, y_i^2) + F(\beta(m_i))u(2, y_i^2) = F(\beta(m_i))v(y_i^2)$.

Therefore, System 2's problem is to choose m_i such that it solves:

$$\max_{m_1,\ldots,m_K} \sum_{i=1}^{K} F(\beta(m_i))v(y_i^2) \quad s.t. \quad m_i \geq 0 \text{ for all } i, \text{ and } \frac{1}{K}\sum_{i=1}^{K} m_i \leq B.$$

Let $\hat{F}(m) = F(\beta(m))$. Because of Assumption 3 on beliefs, the objective function is concave, and the first order conditions are sufficient for optimality:

$$\hat{f}(m_i) \cdot v(y_i^2) = \lambda - \mu_i, \qquad \text{and hence} \quad m_i = \hat{f}^{-1}\left(\frac{\lambda - \mu_i}{v(y_i^2)}\right), \tag{1}$$

where $\hat{f}$ is the derivative of $\hat{F}$, μ_i is the Lagrange multiplier of the first constraint, and λ is the Lagrange multiplier of the second constraint (since $\hat{F}$ is concave, $\hat{f}$ is injective and its inverse is well defined).

2.2 Principal's problem and main result

The Principal's problem consists on choosing utility function $v(y^2)$ for System 2, as well as a budget B, in order to maximize genetic fitness given System 2's induced behavior. The Principal has two tools to influence the behavior of System 2: by choosing $v(y^2)$, the Principal gives System 2 incentives to consider a certain state more important than other; by choosing B, the Principal decides how much to limit the autonomy of System 2. It turns out that these two tools are enough for the Principal to achieve optimal fitness asymptotically (as it is standard in contract theory, the optimal outcome is defined as the maximal payoff for the Principal *in the absence of constraints*; in this case, it is the maximal fitness the Principal could attain if it observed y_i^2 and chose a_i directly).

Proposition 1. *There exists an asymptotic equilibrium, which is characterized by:*

$$v(y^2) = \frac{1}{f(y^2)}, \qquad \textit{and} \qquad B = \mathbb{E}[y^2],$$

for the case $\beta(m) = m$. Moreover, this asymptotic equilibrium is unique, up to outcome-irrelevant transformations of System 1's beliefs.[15]

The intuition for the result is as follows. Because the Principal cannot include y^1 into System 2's utility function, it chooses a utility function $v(y^2)$ for System 2 which is biased towards always choosing action $a = 2$, and the more so the higher y^2 is. Note that the Principal could have an unbiased System 2, by choosing $v(y^2) = u(2, y^2) - u(1, y^2) = y^2 - \mathbb{E}[y^1]$: System 2 would be unbiased, but the Principal would be unable to extract System 2's information about y^2. Thus, by biasing System 2 through its utility function, and then disciplining it through the budget B, with a limited amount of self-control, so that System 2 is constrained in its autonomy, the Principal can achieve optimal fitness. It is the combination of being able to choose both $v(y^2)$ and B that allows the Principal to reach optimal fitness asymptotically.[16]

Observe that $v(y^2)$ is increasing in y^2 (as $f(y^2)$ is strictly decreasing by virtue of $F(y^2)$ being strictly concave in y^2), so System 2 gives more importance to $a = 2$ being implemented in states in which y^2 is higher. In a sense, the Principal makes System 2 an "advocate" for action $a = 2$ (as in the case of Che and Kartik, 2009, in which the Principal purposefully chooses a biased Agent, to generate incentives for the Agent to collect information in order to convince the Principal). Because the Principal disciplines System 2 by limiting the amount of messages it can send, System 2 has an incentive to send high m_i only when y_i^2 is truly high, following the intuition of the literature in decision linkage (Jackson and Sonnenschein, 2007; Frankel, 2014). In fact, this is analogous to the example in the Introduction about the principal of a school choosing an incentive scheme that encourages teachers to inflate grades, but who sets a cap on the number

[15] In particular, for any beliefs $\beta(m)$, there is an asymptotic equilibrium with $v(y^2) = \frac{1}{\tilde{f}(\beta^{-1}(y^2))}$ and $B = \mathbb{E}[\beta^{-1}(y^2)]$, that yields the exact same outcome as the proposed asymptotic equilibrium. Note that, although the budget seems arbitrary, it is not so. The budget is uniquely determined, up to "changes in scale": if System 1 would interpret every message m as $\beta(m) = m/2$, then the Principal would achieve the same final outcome by endowing System 2 with a budget of $2B$. In the remainder of the paper, I will assume $\beta(m) = m$, as it simplifies all the expressions without affecting any results.

[16] Note that the Principal can achieve the first best asymptotically even if System 2 had distorted beliefs. If System 2 believed that $y^1 \sim \tilde{F}(y^1)$, all the previous arguments hold if the Principal chose $v(y^2) = \frac{1}{\tilde{f}(y^2)}$. The only requirement is that, in that case, the distorted beliefs operate in an evolutionarily long time (an assumption which is made for example by Rayo and Robson, 2016).

of students who can receive a grade of "A", while allowing the teacher to choose which students get each grade (Frankel, 2014). In conclusion, Proposition 1 shows that, *against the popular intuition that "more self-control is always more adaptive", there is actually an optimal amount of self control that maximizes genetic fitness, thus explaining why humans evolved a limited (rather than perfect) self-control.*

2.3 Evidence from neuroscience

I attempt here to provide the intuition (based on the evidence we have from neuroscience about the functioning of the brain) for the assumptions of the model.

I begin with Assumption 1, regarding the information that each player has. As I discussed above, y^2 refers to the particular social norms of the individual's group about cheating, that are extremely specific and too fast-changing to be encoded directly by the genes (Richerson and Boyd, 2005; Henrich *et al.*, 2001), and hence unobservable by the Principal (and also by System 1 since it is assumed that its behavior is genetically determined), although easily observable by System 2. However, while the particular social norm about cheating cannot be encoded in the genes, the *value of following the social norms* is stable enough that it can, hence having y^2 as an argument in System 2's utility function. In other words, while the genes cannot encode the particular social norms for each society, they can encode the disutility of breaking whichever social norm happen to be in place. On the flip side, y^1 is not observable for System 2, nor can it be included in its utility function i.e. it represents characteristics for which the individual is not aware ("bottom up signals", such as hormones). This makes sense if we think of y^1 as the fitness value of a cheating opportunity: for example in the case of sexual cheating, this value can depend on inferring the health and reproductive status of the potential sexual partner from subtle cues. In primates, the orbitofrontal cortex (OFC) has been showed to respond to social status and sexual cues (Watson and Platt, 2012), and the OFC can affect the value signal for y^1 in the vmPFC which, together with the self-control exerted by the dlPFC (i.e. System 2's message m) are compared by System 1.[17]

Support for the last part of Assumption 1, regarding the impossibility of including all relevant variables in System 2's utility function, stems from the *modularity* of the brain. There is ample evidence that the brain and the mind are modular (they are composed of several more-or-less insulated modules which do not necessarily share information); and behavior stems from a conflict between competing modules (Fodor, 1983; Tooby and Cosmides, 1992; Cosmides and Tooby, 1994; Pinker, 1997; Barrett and Kurzban, 2006; Sanfey *et al.*, 2006; Kurzban and Athena Aktipis, 2007). Self-control can therefore be understood as a manifestation of a conflict between brain systems, an approach that has been used previously (Thaler and Shefrin, 1981; Fudenberg and Levine, 2006; Brocas and Carrillo, 2014). Modularity might be especially important in humans as compared to other species, because as our brain became larger in the course of evolution, the cost of connecting distant regions grew, and so there was a change in the pattern of connectivity, that exacerbated local connections and specialization (Gazzaniga, 2012). From a computational point of view, a build of the brain composed of a number of conflicting modules might be the optimal design to solve complex tasks (Livnat and Pippenger, 2006).[18] Behavior would then result from the aggregation of these brain modules, and these modules can recommend different behaviors, and hence be in conflict with each other.[19] Further evidence of modularity and encapsulation comes from studies that show that people do not have conscious access to their own mental states and processes (Gazzaniga, 1970; Nisbett and Wilson, 1977; Libet *et al.*, 1983). This assumption has been incorporated into economics by Brocas and Carrillo (2008), who consider that there is a problem of asymmetric information between different brain subsystems (see also Brocas and

[17] I thank a referee for suggesting that these neural circuits would affect the value signal before it reaches System 1, i.e. dmPFC and IPS regions.

[18] According to Sanfey *et al.* (2006): "There is a long legacy of research within psychology, strongly supported by findings from neuroscience, to suggest that human behavior is not the product of a single process, but rather reflects the interaction of different specialized subsystems. Although most of the time these systems interact synergistically to determine behavior, at times they compete, producing different dispositions towards the same information". (Ainslie, 2001, p.43) agrees with this view. See also McClure *et al.* (2004), and Haidt (2006) for a book-length treatment.

[19] Bisin and Iantchev (2016) showed a similar result, namely that a hierarchical mind with conflicting modules is evolutionarily adaptive (see also Kurzban, 2012). I would like to emphasize that *I am not arguing that there exactly two subsystems in the mind*, but rather that there are a number of subsystems that exhibit modularity, and the framework with two subsystems serves as a useful modeling device (Alós-Ferrer and Strack, 2014).

Carrillo, 2014, for a review of the implications of modularity in economic applications).

I turn now to justify the inclusion of System 1 in the model, as well as Assumption 2. An area in the brain, namely the ventromedial prefrontal cortex (vmPFC), is responsible for valuation of different alternatives (Kable, 2013), and exercising self-control happens by activating another area, the dorsolateral prefrontal cortex (dlPFC), which modulates the value signal in the vmPFC (for example, a dieter can reduce the value of a tasty but unhealthy meal by exerting self-control, what activates their dlPFC and reduces that value of the meal in the vmPFC, Hare *et al.*, 2009).[20] This final value is passed on to regions specialized in comparing the values (dmPFC and IPS), that then choose one course of action (Hare *et al.*, 2011). System 2 is therefore assumed to represent the dlPFC, exerting self-control in order to modulate the value of the alternatives. This is consistent with the fact that the dlPFC is responsible for executive function (that involves task-switching, planning, and working memory, Elliott, 2003), as well as directly on the exertion of self-control (Knoch *et al.*, 2006; Hare *et al.*, 2009; Figner *et al.*, 2010). System 1 represents the comparator regions (dmPFC and IPS), that simply take the action that has the highest value (in terms of genetic fitness), and implement such action. Further evidence has showed that disruption of the left dlPFC leads subjects to choose immediate smaller options over delayed larger ones (Figner *et al.*, 2010), and disruption of the right dlPFC makes subjects more likely to accept unfair (but "tempting") offers in the Ultimatum Game (Knoch *et al.*, 2006). Moreover, patients with lesions in their dlPFC behaved more dishonestly in order to increase their payoffs (Zhu *et al.*, 2014). This evidence strongly suggests a role for the dlPFC in self-control.

Assumption 3 is a technical assumption for the maximization problem to be well defined. Finally, the justification for Assumption 4 stems from evidence that humans have a limited amount of self-control, a phenomenon known as *ego depletion* (Baumeister *et al.*, 1998; Muraven *et al.*, 1998); see Achtziger *et al.* (2016) for an application in economics. In particular, the ego depletion literature makes the following two claims: 1) people have a finite amount of willpower which becomes depleted as they exercise self-control (Baumeister *et al.*, 1998; Muraven *et al.*,

[20] Notably, when the choice task is reframed in such a way that temptation is reduced, the choice that previously required self-control can be taken without exertion of self-control (i.e. activation of the dlPFC Magen *et al.*, 2014).

1998); and 2) all tasks that require self-control use the same "stock of willpower" (Baumeister and Tierney, 2011). However, there has been a recent controversy regarding the ego depletion literature: while a 2010 meta-analysis found this effect to be robust (Hagger, 2010), a more recent meta-analysis which used a methodology that accounted for small sample experiments and publication bias, found a zero effect (Carter and McCullough, 2014), and several replication attempts have failed to find an effect distinguishable from zero (Alós-Ferrer *et al.*, 2019), including a multi-lab replication of a classic version of the experiment used to determine the ego depletion phenomenon (Hagger and Chatzisarantis, 2016). However, another recent replication attempt did find a small but significant effect (Dang *et al.*, 2021). As of today, the debate is still open and it seems that the ego depletion effect, if confirmed to exist, is probably heavily context-dependent. Because of that, the framework provided in this paper can potentially shed light on the nature of the ego depletion effect, and self-control more generally.

Despite the controversy in the ego depletion literature, Assumption 4 is the least in need of external justification because, given the rest of the model, assuming that the Principal can choose B (in addition to System 2's utility function) is a natural assumption, and in line with the spirit of the literature on the evolution of preferences discussed in the Introduction. Moreover, this paper can add to the ongoing controversy by providing a theoretical framework that yields testable predictions (Friese *et al.*, 2019).

3 Is self-control like a muscle? An imperfect metaphor

In the following sections I apply the model to several interesting facts found in the empirical literature on self-control. Before proceeding, I would like to point out that the experimental tasks used to measure self-control, such as not eating a marshmallow or keeping one's hands in cold water (Mischel *et al.*, 1972; Baumeister *et al.*, 2007), can seem far from the stylized example of cheating vs. breaking social norms. However, we must keep in mind that the economic environment in which humans evolved was rather poor (with limited economic choices), and thus why it can be appropriately modeled by a series of K binary choices between cheating and respecting a social norm. In today's world, the number of economic choices is extremely large for any individual, and that is why many tasks can be used to measure self-control, as long as they generate a tradeoff in the individual between, for example, following the instruction of the experimenter vs. cheating.[21]

3.1 Self-control becomes exhausted

Researchers of self-control have found that, after an individual exerts self-control in a given task, they show less self-control in subsequent (possibly unrelated) tasks, a phenomenon which has been termed *ego depletion* (Baumeister *et al.*, 2007, but see Section 2.3 above for a nuanced view of this phenomenon in light of the recent replication crisis). Note that as outside observers, we cannot directly observe m_i, only whether $a_i = 1$ or $a_i = 2$. In other words, if we ask a subject to keep their hands in ice-cold water for one minute, we can only observe whether they succeeded at doing so or not; we cannot observe the amount of self-control they exerted. Because of that, I define a measure of self-control which simply amounts to counting the times when $a_i = 2$, in all tasks except the first.

Definition 3. *Reveled self-control in subsequent tasks is given by* $\sum_{i=2}^{K} \mathbb{1}\{a_i = 2\}$.

Suppose that System 2 is given a task in two different treatments τ: let $\tau = C$ be a control condition, and $\tau = S$ be a treatment that requires self-control. Crucially, the first task's fitness $y_1^1(\tau)$ is such that $y_1^1(\tau) = \rho(\tau) + y_1^1$, where $y_1^1 \sim F$. In the control $\tau = C$, we have that $\rho(C) = 0$, so that we are back to the original model. However, in the treatment, we have $\rho(S) > 0$. All other tasks $i \geq 2$ are such that $y_i^1(\tau) = y_i^1$ as usual. The interpretation of $\rho(S)$ is that in the treatment condition, the temptation is higher for the first task. The following Proposition formalizes the ego depletion effect.

[21] This is also the reason why limited self-control seems to be at the root of several modern behaviors, such as alcohol consumption, smoking, not saving or exercising enough, etc. (Baumeister *et al.*, 1994). I discuss the implications of this *evolutionary mismatch* in Section 4.2.

Proposition 2. *Revealed self-control in subsequent tasks is smaller in the Treatment versus the Control condition.*

The intuition for Proposition 2 is simple: as $\rho(S)$ increases, the amount of self-control exerted in the first task increases too, and the remaining budget $B - \frac{m_1}{K}$ decreases accordingly. Since there is less budget left in the Treatment than in the Control condition, then System 2 is forced to send smaller messages, and therefore to have smaller revealed self-control. This shows that ego-depletion is a natural consequence of the fact that System 2 has a limited budget of self-control and, in that respect, the muscle metaphor is apt. We turn now to whether the "self-control muscle" can become stronger.

3.2 Self-control grows over time when it is needed

For at least twenty-five centuries, philosophers have been advocating for people to exercise self-control, in order to grow it over time. For example, the philosopher Epictetus recommended to his students to strengthen their self-control daily. Recently, it has been showed that this is a good advice, and that self-control can grow when necessary, just as a muscle. For example, exercising self-control in any domain (be it physical exercise, dutifully attending one's academic obligations, or improving personal finance) improved people's performance in an unrelated self-control task (Muraven *et al.*, 1999; Oaten and Cheng, 2006b,a, 2007). As I show next, having self-control depend on previous needs is a simple way to adapt to a changing environment.

So far I have assumed that the distribution of y^2 is fixed over time. In this section, I assume instead that the distribution of y^2 depends on a parameter θ, which is unknown *a priori*: y^2 is distributed according to $G(y^2|\theta) \sim \mathcal{N}_+(\theta, \sigma^2)$, where $\mathcal{N}_+$ refers to a truncated normal (that takes only non-negative values). Suppose that the Principal-Agent relation happens at different time periods $t \in \{0, 1, 2, \dots\}$. In this extension of the model, the Principal can design a budget $B(t)$ that updates at the end of each period t. Let $\bar{y}^2(t-1) = \frac{1}{K}\sum_{i=1}^{K} y_i^2(t-1)$ be the average level of self-control "needed" in period $t-1$. I assume that the updating process of B_t can be conditional on $\bar{y}^2(t-1)$.

Proposition 3. *The self-control budget $B(t)$ is increasing in $\bar{y}^2(t-1)$.*

The main intuition behind Proposition 3 is that $\bar{y}^2(t-1)$ is informative about θ, and therefore the Principal tailors the level of the self-control budget to the distribution of y^2, which depends on θ.[22] The Principal anticipates that high realizations of $\bar{y}^2(t-1)$ imply that θ must be higher than expected, in which case there will be higher realizations of $y^2(t)$, and thus the optimal self-control budget should be higher.

So far we have seen that self-control behaves much like a muscle: it gets tired when we use it continuously (Proposition 2), and it grow over time (Proposition 3). However, the muscle metaphor might be imperfect: self-control, unlike a muscle, might not need to depend directly on a physical resource such as glucose.

3.3 Is glucose consumed or only monitored?

In the last decade, a series of studies have found that self-control is linked to glucose (for example, low blood sugar is correlated with unlawful behavior, Gailliot and Baumeister, 2007). One proposed hypothesis is that the brain consumes extra glucose during self-control tasks (Gailliot *et al.*, 2007). However, Kurzban (2010) re-analyzed the data from Gailliot *et al.* (2007) and found no support for the original hypothesis, proposing instead that glucose should be considered as a variable that is monitored, rather than as an input that is consumed, into the decision-making process: I call this the **glucose monitoring hypothesis**.

Long-term investments are worthless (in terms of genetic fitness) if the individual dies and hence cannot reap the benefits of the investment. The question then becomes: how does the Principal determine the relative survival prospects of the individual? I assume that the level of glucose γ can be monitored, and that this can serve as an input to assess potential fitness.[23] Let

[22] The assumption that $G(y^2|\theta)$ is distributed as a truncated normal is not innocuous: the normal distribution has the Monotone Likelihood Ration Property, which implies that higher realizations of y^2 imply that a higher θ is more likely.

[23] Low blood glucose starts a cascade of hormonal and neurophysiological changes, and glucose and energy balance (*homeostais*) are crucial for the survival and thriving of the individual (Williams and Elmquist, 2012). Therefore, it is plausible that low blood glucose affects self-control indirectly, through neuroendocrine changes.

γ be a function that transforms the monitored level of glucose into "fitness units". I assume that y^1 does not change with glucose, but that y^2 is obtained from a distribution $G(y^2|\gamma)$, and that the budget $B(\gamma)$ can depend on γ.

Proposition 4. *There is a unique asymptotic equilibrium with* $v(y^2) = \frac{1}{f(y^2)}$ *and* $B(\gamma) = \mathbb{E}[y^2|\gamma]$, *for beliefs* $\beta(m) = m$ *(up to outcome irrelevant-transformations of System 1's beliefs).*

Note that the budget $B(\gamma)$ is given by the expectation of y^2 conditional on γ: even though glucose is not consumed, it is used as an indicator of how much self-control is optimal to endow System 2 with. Proposition 4 therefore offers a plausible mechanism as for how glucose could serve as a monitored variable, rather than an input that is consumed.

4 Welfare implications

The model presented in this paper has strong normative implications. The first, and most important, is the idea that individuals might be choosing actions that are against their own interest. This in turn implies that individuals are subject to internalities (Berridge, 2003; Berridge and O'Doherty, 2014), what has given birth to the recent field of behavioral welfare economics.[24]

4.1 Decoupling choice from welfare

A recent literature is taking seriously the idea that revealed preferences do not always reflect the welfare of the individual (Kahneman *et al.*, 1997; Kahneman and Sugden, 2005; Allcott and Taubinsky, 2015; Chetty, 2015). Suppose that the social planner aims to maximize a weighted average of the utilities from System 2 and the Principal/System 1, $\mathbb{E}\left[\zeta \cdot u(a, y^2) + (1-\zeta) \cdot y^a\right]$ (note that Principal and System 1 have the same utility function, i.e. genetic fitness y^a, and so the social planner only needs to consider the relative weights of utilities between System 2 and Principal/System 1). Then, it is easy to see that whenever the weight given to System 2

[24] Chetty (2008); Mullainathan *et al.* (2012); Allcott *et al.* (2014); Chetty (2015); Allcott and Taubinsky (2015); Taubinsky and Rees-Jones (2018); Allcott *et al.* (2019); Bernheim and Taubinsky (2019); Jimenez-Gomez (2017), among others.

is positive ($\zeta > 0$), the social planner would like to endow System 2 with a larger self-control budget.

Proposition 5. *The socially optimal budget B^* is larger than the budget B provided by natural selection.*

What Proposition 5 amounts to, is that the budget generated by natural selection is too small as long as the social planner cares about System 2 at all. The reason is that when the budget is $B = \mathbb{E}[y^2]$ (the one provided by natural selection), this will result in the maximization of y^a. That means that action $a = 1$ is taken too often, and therefore the maximization of the convex combination of $u(a, y^2)$ and y^a would require an increase in the probability of action $a = 2$ being chosen, what is achieved with a larger self-control budget.[25]

The implication is that the social planner should engage in policies that either increase people's level of self-control, or reduce the availability of temptations, and this is consistent with (for example) evidence that suggests smokers are happier when cigarettes are taxed (Gruber and Mullainathan, 2006).

4.2 Evolutionary mismatch

While the example I have used so far to discuss self-control (and lack thereof) is breaking social norms, today self-control plays a large role in a myriad of different behaviors. Lack of self-control has been linked to consumption of alcohol, cigarettes, and illicit drugs, undersaving, lack of exercise, unhealthy eating, etc. (Heatherton and Wagner, 2011). This is validated by research that shows that unhealthy eating is associated with neural activity consistent with self-control, such as decreased activity and connectivity in the dorsolateral prefrontal cortex (Rhodes *et al.*, 2013, see also Section 2.3). It has been argued that humans suffer an **evolutionary**

[25] This reasoning is true for any social welfare function (not just the sum of utilities) that has System 2's utility as an argument, and is increasing on that argument. It has been suggested to me that I might want to consider fitness maximization as a welfare criterion. However, there is no normative reason why we should care about fitness maximization. If we did, this would lead to some paradoxical consequences, such that we should attempt to maximize the number of children people have (which is what fitness measures) over their own happiness.

mismatch: human behavior was optimized by natural selection for a hunter-gatherer lifestyle but, because our environment has changed rapidly, those same behaviors are not adaptive in the current environment (Spinella, 2003). Because of this evolutionary mismatch, we live in a period where arguably there are more temptations than ever before. Addictive drugs such as alcohol and tobacco can be legally purchased by adults over a certain age; sugar and red meat (which were scarce in our evolutionary environment) are now readily available; and lack of exercise and being overweight has been linked to an increase in the risk of contracting certain diseases such as cancer (Kushi *et al.*, 2012).[26] In order to capture this increased availability of temptations, suppose that there is a shift in the distribution of y^1, so that the current value is $\hat{y}^1 = \rho + y^1$, where $\rho > 0$ is a constant and y^1 is the original value. This change happens too fast in the evolutionary time scale for the Principal and System 1 to react to it, so that v and B are still given by $v(y^2) = \frac{1}{f(y^2)}$ and $B = \mathbb{E}[y^2]$. I am agnostic about whether this change is stable enough that System 2 perceives it, in which case I say it is *sophisticated*, or whether System 2 is not aware about this change in the distribution, in which case I say it is *naive*. It turns out that in both cases System 2 will fail to resist temptation more often in the current environment.

Proposition 6. *Irrespectively of whether System 2 is sophisticated or naive, System 1 chooses the "impulsive" action $a = 1$ more often in the current environment than in the original environment.*

The intuition behind Proposition 6 is as follows: when y^1 increases to $\hat{y}^1$, System 2 needs to exert more self-control *in each task* in order to have $a = 2$ implemented. Because the budget is fixed, that means that it is now "more expensive" to have $a = 2$ implemented for each possible realization of y^2, and therefore the probability that $a = 1$ increases across the board. The reader might argue that if the distribution of y^1 has changed over time, then the Principal should have changed B as well. However, gene evolution is (usually) a slow process: humans' genes have not changed fast enough to adapt to the temptations (in the form of alcohol, tobacco and other drugs, fatty and sugary foods, etc., Buss, 2015), which have increased exponentially. Because of that, it

[26] Unhealthy eating is a primary factor in the rise of obesity across the world (Crino *et al.*, 2015). Moreover, non-communicable diseases (such as stroke and heart attack, cancers, chronic lung diseases, etc.) accounted for more than 60% of global deaths in 2008 (Rünger and Wood, 2015).

seems reasonable to consider that the change in environment $\hat{y}^1 = y^1 + \rho$ happened faster than our genes could evolve, and therefore that B has remained fixed. For the same reason, System 1, who shares the same utility function as the Principal, and could be thought of as "hard-wired" by the Principal, does not adapt to the change in the environment. Proposition 6 implies that whatever conclusions can be drawn from the present model will be exacerbated in the current environment, because of an increase in y^1: lack of self-control decreases individual welfare in the original environment (under y^1, Proposition 5), but Proposition 6 shows that it will decrease individual welfare *even more* due to evolutionary mismatch in the current environment (under $\hat{y}^1$). Because of evolutionary mismatch, we are at a point where System 2 is receiving too little self-control budget, to cope with all the temptations of the modern world (and firms might be exploiting this fact Akerlof and Shiller, 2015; Jimenez-Gomez, 2017).

5 Robustness: Dynamic choice

So far, the problem we have analyzed had System 2 observing $(y_i^2)_{i=1}^K$, and then making the choice of m_i. However in real life decisions do not happen simultaneously, but rather as they appear into a person's life. In this section I show that, asymptotically, the solution for the static and the dynamic problems coincide (I follow Jackson and Sonnenschein (2007) in solving the static problem first, and then showing that the solutions for the static and dynamic problem are asymptotically identical). Suppose that the Principal chooses the same utility function $v(y^2) = \frac{1}{f(y^2)}$ and budget $B = \mathbb{E}[y^2]$ which solved the static problem in Proposition 1. Moreover, suppose System 2 naively chooses $m_i = y_i^2$ as long as that option is within the budget, and otherwise chooses the maximal m_i possible. In other words:

$$m_i = \begin{cases} y_i^2 & \text{if } \frac{1}{K}\sum_{j=1}^{i} y_j^2 \leq B, \\ \max\left\{B \cdot K - \sum_{j=1}^{i-1} y_j^2, 0\right\} & \text{if } \frac{1}{K}\sum_{j=1}^{i} y_j^2 > B. \end{cases} \tag{2}$$

In that case, System 2's average payoff is given by $\frac{1}{K}\{\sum_{i=1}^K u(1, y_i^2) + F(m_i) \cdot v(y_i^2)\}$, which converges to $\mathbb{E}[u(1, y^2) + F(y^2) \cdot v(y^2)]$. That means that the average payoff in the dynamic case converges to the payoff in the static case. We have the following result.

Proposition 7. *For every $\nu > 0$ there exists $\bar{K}$ such that for all $K > \bar{K}$, if System 2 chooses m_i as in Equation 2, it obtains a payoff that is smaller than that of the static case by at most ν.*

Proposition 7 implies that the dynamic behavior of System 2 described in Equation 2 approximates the optimal behavior for System 2 in the static case. Thus, we can consider that the static baseline model, as well as its extensions presented up to now, are useful approximations of a dynamic model in which Systems 1 & 2 face decisions (of actions a_i and messages m_i) sequentially.

6 Conclusion

Self-control problems are at the heart of much human suffering, and are pervasive. Since the human brain and mind are the result of an evolutionary process that selected for the most adapted individuals, this begs the question of why did humans not evolve a stronger, or even perfect, self-control (Hayden, 2019). I have provided an answer: the current (and limited) self-control that humans posses evolved precisely to make humans behave against their own interest (as judged by their evolved utility function), in instances in which that behavior would be genetically adaptive. This is the solution to an endogenous conflict of interest between the genes and the human individual; a conflict that arises from the impossibility of the genes to include all the relevant variables into the individual's (specifically, System 2's) utility function.

The model presented in this paper can explain several stylized facts found in the empirical literature on self-control, such as the fact that self-control becomes depleted with usage in the short-run, and also that it can grow over time. Moreover, even if self-control becomes depleted when glucose levels are low, the model shows that glucose is not necessarily consumed to exert self-control. This paper can also help shed light on the ongoing debate about whether the results from the ego depletion literature are real (a controversy I described in Section 2.3). It has been argued that in order to make progress in settling this debate, better theoretical models of self-control are needed (Carter and McCullough, 2014; Friese *et al.*, 2019), and this paper could provide a theoretical framework with which to confront the conflicting empirical evidence of ego-depletion effects, while also providing a critical view on some of their findings (as in the case of the impact of glucose on self-control). The model is also helpful in thinking about normative economics. While self-control problems have most likely always been detrimental for welfare (Proposition 5), this phenomenon has been exacerbated in recent decades by an abundance of temptations (Proposition 6), and there is ample evidence that in the present day self-control problems are an important source of welfare loss (Baumeister *et al.*, 1994).

The framework presented in this paper represents a step towards connecting research in economics with current topics in cognitive science and neuroscience, and towards understanding the evolutionary origin of human nature, one of the most fascinating questions we can ask, with profound implications for cognitive science, positive and normative economics, and policy.

References

Achtziger, A., Alós-Ferrer, C. and Wagner, A. K. (2016) The impact of self-control depletion on social preferences in the ultimatum game, *Journal of Economic Psychology*, **53**, 1–16.

Aghion, P. and Tirole, J. (1997) Formal and Real Authority in Organizations, *The Journal of Political Economy*, **105**, 1–29.

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

Akerlof, G. A. and Shiller, R. J. (2015) *Phishing for Phools: The Economics of Manipulation and Deception*, Princeton University Press.

Alan, S., Boneva, T. and Ertac, S. (2019) Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit, *The Quarterly Journal of Economics*, **134**, 1121–1162.

Allcott, H. (2011) Social norms and energy conservation, *Journal of Public Economics*, **95**, 1082–1095.

Allcott, H., Lockwood, B. B. and Taubinsky, D. (2019) Should we tax sugar-sweetened beverages? An overview of theory and evidence, *Journal of Economic Perspectives*, **33**, 202–227.

Allcott, H., Mullainathan, S. and Taubinsky, D. (2014) Energy policy with externalities and internalities, *Journal of Public Economics*, **112**, 72–88.

Allcott, H. and Taubinsky, D. (2015) Evaluating behaviorally-motivated policy: Experimental evidence from the lightbulb market, *American Economic Review*, **105**, 2501–38.

Alonso, R., Brocas, I. and Carrillo, J. (2013) Resource allocation in the brain, *Review of Economic Studies*.

Alonso, R. and Matouschek, N. (2008) Optimal delegation, *Review of Economic Studies*, **75**, 259–293.

Alós-Ferrer, C., Ritschel, A., Garcia-Segarra, J. and Achtziger, A. (2019) Habituation does not rescue depletion: Two tests of the ego-depletion effect, *Journal of Economic Psychology*, **75**, 102122.

Alós-Ferrer, C. and Strack, F. (2014) From dual processes to multiple selves: Implications for economic behavior, *Journal of Economic Psychology*, **41**, 1–11.

Barrett, H. C. and Kurzban, R. (2006) Modularity in cognition: framing the debate., *Psychological review*, **113**, 628.

Baumeister, R. F., Bratslavsky, E., Muraven, M. and Tice, D. M. (1998) Ego depletion: is the active self a limited resource?, *Journal of Personality and Social Psychology*, **74**, 1252.

Baumeister, R. F., Heatherton, T. F. and Tice, D. M. (1994) *Losing control: How and why people fail at self-regulation.*, Academic Press.

Baumeister, R. F. and Tierney, J. (2011) *Willpower: Why Self-Control is the Secret to Success*, Penguin Group US.

Baumeister, R. F., Vohs, K. D. and Tice, D. M. (2007) The Strength Model of Self-Control, *Current Directions in Psychological Science*, **16**, 351–355.

Benhabib, J. and Bisin, A. (2005) Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions, *Games and Economic Behavior*, **52**, 460–492.

Bernheim, B. D. and Rangel, A. (2004) Addiction and cue-triggered decision processes, *American Economic Review*, pp. 1558–1590.

Bernheim, B. D. and Taubinsky, D. (2019) *Behavioral Public Economics*, vol. 1, Elsevier B.V.

Berridge, K. (2003) Irrational pursuits: Hyper-incentives from a visceral brain, *The psychology of economic decisions*, **1**, 17–40.

Berridge, K. C. and O'Doherty, J. P. (2014) From experienced utility to decision utility, *Neuroeconomics*, pp. 335–351.

Bisin, A. and Iantchev, E. (2016) Evolutionary Selection of Modular Decision Architectures, *Working Paper*.

Blattman, C., Jamison, J. C. and Sheridan, M. (2017) Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia, *American Economic Review*, **107**, 1165–1206.

Boyd, R. and Richerson, P. J. (1985) *Culture and the evolutionary process*, University of Chicago Press.

Brocas, I. and Carrillo, J. D. (2008) The brain as a hierarchical organization, *The American Economic Review*, **98**, 1312–1346.

Brocas, I. and Carrillo, J. D. (2014) Dual-process theories of decision-making: A selective survey, *Journal of Economic Psychology*, **41**, 45–54.

Brocas, I. and Carrillo, J. D. (2021) Value computation and modulation: A neuroeconomic theory of self-control as constrained optimization, *Journal of Economic Theory*, **198**, 105366.

Buss, D. M. (2015) *Evolutionary Psychology: The New Science of the Mind, Fifth Edition*, Taylor & Francis.

Carter, E. C. and McCullough, M. E. (2014) Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated?, *Frontiers in Psychology*, **5**, 1–11.

Cerigioni, F. (2021) Dual decision processes: retrieving preferences when some choices are automatic, *Journal of Political Economy*, **129**, 1667–1704.

Che, Y.-K. and Kartik, N. (2009) Opinions as incentives, *Journal of Political Economy*, **117**, 815.

Chetty, R. (2008) Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods, *Annual Review of Economics*, **1**, 451–488.

Chetty, R. (2015) Behavioral Economics and Public Policy: A Pragmatic Perspective, *American Economic Review*, **105**, 1–33.

Cosmides, L. (1989) The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task, *Cognition*, **31**, 187–276.

Cosmides, L. and Tooby, J. (1992) Cognitive adaptations for social exchange, *The adapted mind: Evolutionary psychology and the generation of culture*, **163**, 163–228.

Cosmides, L. and Tooby, J. (1994) Origins of domain specificity: The evolution of functional organization, *Mapping the mind: Domain specificity in cognition and culture*, pp. 85–116.

Crino, M., Sacks, G., Vandevijvere, S., Swinburn, B. and Neal, B. (2015) The Influence on Population Weight Gain and Obesity of the Macronutrient Composition and Energy Density of the Food Supply, *Current obesity reports*, **4**, 1–10.

Dang, J., Barker, P., Baumert, A., Bentvelzen, M., Berkman, E., Buchholz, N., Buczny, J., Chen, Z., De Cristofaro, V., de Vries, L., Dewitte, S., Giacomantonio, M., Gong, R., Homan, M., Imhoff, R., Ismail, I., Jia, L., Kubiak, T., Lange, F., Li, D. Y., Livingston, J., Ludwig, R., Panno, A., Pearman, J., Rassi, N., Schiöth, H. B., Schmitt, M., Sevincer, A. T., Shi, J., Stamos, A., Tan, Y. C., Wenzel, M., Zerhouni, O., Zhang, L. W., Zhang, Y. J. and Zinkernagel, A. (2021) A Multilab Replication of the Ego Depletion Effect, *Social Psychological and Personality Science*, **12**, 14–24.

Dasgupta, P. and Maskin, E. (2005) Uncertainty and Hyperbolic Discounting, *The American Economic Review*, **95**, 1290–1299.

Duckworth, A. L. and Seligman, M. E. (2005) Self-discipline outdoes IQ in predicting academic performance of adolescents, *Psychological science*, **16**, 939–944.

Elliott, R. (2003) Executive functions and their disorders: Imaging in clinical neuroscience, *British medical bulletin*, **65**, 49–59.

Epictetus (2008) *Discourses and Selected Writings*, Penguin classics, Penguin Books Limited.

Fawcett, T. W., McNamara, J. M. and Houston, A. I. (2012) When is it adaptive to be patient? A general framework for evaluating delayed rewards, *Behavioural Processes*, **89**, 128–136.

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E. and Weber, E. U. (2010) Lateral prefrontal cortex and self-control in intertemporal choice., *Nature neuroscience*, **13**, 538–539.

Fodor, J. A. (1983) *The Modularity of Mind*, MIT press.

Frankel, A. (2014) Aligned delegation, *The American Economic Review*, **104**, 66–83.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J. and Inzlicht, M. (2019) Is ego depletion real? An analysis of arguments, *Personality and Social Psychology Review*, **23**, 107–131.

Fudenberg, D. and Levine, D. K. (2006) A dual-self model of impulse control, *The American Economic Review*, pp. 1449–1476.

Gailliot, M. T. and Baumeister, R. F. (2007) The physiology of willpower: Linking blood glucose to self-control, *Personality and Social Psychology Review*, **11**, 303–327.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E. and Schmeichel, B. J. (2007) Self-control relies on glucose as a limited energy source: willpower is more than a metaphor., *Journal of Personality and Social Psychology*, **92**, 325–36.

Gazzaniga, M. (2012) *Who's in Charge?: Free Will and the Science of the Brain*, Hachette UK.

Gazzaniga, M. S. (1970) *The bisected brain*, vol. 2, Appleton-Century-Crofts.

Gigerenzer, G. and Hug, K. (1992) Domain-specific reasoning: Social contracts, cheating and perspective change, *Cognition*.

Gruber, J. and Mullainathan, S. (2006) Do cigarette taxes make smokers happier?, in *Happiness and Public Policy*, Springer, pp. 109–146.

Hagger, M. S. (2010) Self-regulation: An important construct in health psychology research and practice, *Health Psychology Review*, **4**, 57–65.

Hagger, M. S. and Chatzisarantis, N. L. D. (2016) A Multilab Preregistered Replication of the Ego-Depletion Effect, *Perspectives on Psychological Science*, **11**, 546–573.

Haidt, J. (2006) *The happiness hypothesis: Finding modern truth in ancient wisdom*, Basic Books.

Hare, T. A., Camerer, C. F. and Rangel, A. (2009) Self-control in decision-making involves modulation of the vmPFC valuation system, *Science*, **324**, 646–648.

Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P. and Rangel, A. (2011) Transformation of stimulus value signals into motor commands during simple choice, *Proceedings of the National Academy of Sciences*, p. 201109322.

Hayden, B. Y. (2019) Why has evolution not selected for perfect self-control?, *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Heatherton, T. F. and Wagner, D. D. (2011) Cognitive neuroscience of self-regulation failure, *Trends in Cognitive Sciences*, **15**, 132–139.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. and McElreath, R. (2001) In search of Homo Economicus: Behavioral experiments in 15 small-scale societies, *American Economic Review, Papers & Proceedings*, **91**, 73–78.

Jackson, M. O. and Sonnenschein, H. F. (2007) Overcoming Incentive Constraints by Linking Decisions, *Econometrica*, **75**, 241–257.

Jimenez-Gomez, D. (2017) Nudging and phishing: A theory of behavioral welfare economics.

Kable, J. (2013) Valuation, Intertemporal Choice and Self-Control, in *Neuroeconomics: Decision Making and the Brain*.

Kacelnik, A. (2003) The evolution of patience., *Time and decision: economic and psychological perspectives on intertemporal choice*, pp. 115 – 138.

Kahneman, D. (2011) *Thinking, Fast and Slow*, Macmillan.

Kahneman, D. and Sugden, R. (2005) Experienced utility as a standard of policy evaluation, *Environmental and resource economics*, **32**, 161–181.

Kahneman, D., Wakker, P. P. and Sarin, R. (1997) Back to Bentham? Explorations of Experienced Utility, *The Quarterly Journal of Economics*, **112**, 375–406.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V. and Fehr, E. (2006) Diminishing Reciprocal Fairness by Disrupting the Right Prefrontal Cortex, *Science*, **314**, 829–832.

Kurzban, R. (2010) Does the brain consume additional glucose during self-control tasks?, *Evolutionary Psychology*, **8**.

Kurzban, R. (2012) *Why Everyone (Else) Is a Hypocrite: Evolution and the Modular Mind*, Princeton University Press.

Kurzban, R. and Athena Aktipis, C. (2007) Modularity and the social mind: Are psychologists too self-ish?, *Personality and Social Psychology Review*, **11**, 131–149.

Kushi, L. H., Doyle, C., McCullough, M., Rock, C. L., Demark-Wahnefried, W., Bandera, E. V., Gapstur, S., Patel, A. V., Andrews, K. and Gansler, T. (2012) American Cancer Society guidelines on nutrition and physical activity for cancer prevention, *CA: A Cancer Journal for Clinicians*, **62**, 30–67.

Libet, B., Gleason, C. A., Wright, E. W. and Pearl, D. K. (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act, *Brain*, **106**, 623–642.

Livnat, A. and Pippenger, N. (2006) An Optimal Brain Can Be Composed of Conflicting Agents, *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 3198–3202.

Magen, E., Kim, B., Dweck, C. S., Gross, J. J. and McClure, S. M. (2014) Behavioral and neural correlates of increased self-control in the absence of increased willpower, *Proceedings of the National Academy of Sciences*, **111**, 9786–9791.

McClure, S. M., Laibson, D. I., Loewenstein, G. and Cohen, J. D. (2004) Separate Neural Systems Value Immediate and Delayed Monetary Rewards, *Science*, **306**, 503 –507.

McGonigal, K. (2011) *The Willpower Instinct: How Self-Control Works, Why It Matters, and What You Can Doto Get More of It*, Penguin.

Mischel, W. (2014) *The Marshmallow Test: Mastering Self-Control*, Little, Brown.

Mischel, W., Ebbesen, E. B. and Raskoff Zeiss, A. (1972) Cognitive and attentional mechanisms in delay of gratification., *Journal of Personality and Social Psychology*, **21**, 204.

Mullainathan, S., Schwartzstein, J. and Congdon, W. J. (2012) A Reduced-Form Approach to Behavioral Public Finance, *Annual Review of Economics*, **4**, 511–540.

Muraven, M., Baumeister, R. F. and Tice, D. M. (1999) Longitudinal Improvement of Self-Regulation Through Practice: Building Self-Control Strength Through Repeated Exercise., *Journal of Social Psychology*, **139**, 446–457.

Muraven, M., Tice, D. M. and Baumeister, R. F. (1998) Self-control as a limited resource: Regulatory depletion patterns., *Journal of Personality and Social Psychology*, **74**, 774–789.

Mylovanov, T. (2008) Veto-based delegation, *Journal of Economic Theory*, **138**, 297–307.

Necker, S. (2016) Why do scientists cheat? Insights from behavioral economics, *Review of Social Economy*, **74**, 98–108.

Netzer, N. (2009) Evolution of time preferences and attitudes toward risk, *The American Economic Review*, **99**, 937–955.

Nisbett, R. E. and Wilson, T. D. (1977) Telling more than we can know: verbal reports on mental processes., *Psychological review*, **84**, 231.

Oaten, M. and Cheng, K. (2006a) Improved self-control: The benefits of a regular program of academic study, *Basic and Applied Social Psychology*, **28**, 1–16.

Oaten, M. and Cheng, K. (2006b) Longitudinal gains in self-regulation from regular physical exercise, *British journal of health psychology*, **11**, 717–733.

Oaten, M. and Cheng, K. (2007) Improvements in self-control from financial monitoring, *Journal of Economic Psychology*, **28**, 487–501.

Pinker, S. (1997) *How the Mind Works*, Norton paperbacks, W.W. Norton.

Postlewaite, A. (2011) *Social norms and preferences*, vol. 1, Elsevier B.V.

Rayo, L. and Becker, G. S. (2007) Evolutionary Efficiency and Happiness, *Journal of Political Economy*.

Rayo, L. and Robson, A. (2016) Biology and the Arguments of Utility, *Working Paper*.

Rhodes, G., Morley, G. and Simmons, L. W. (2013) Women can judge sexual unfaithfulness from unfamiliar men's faces, *Biology Letters*, **9**.

Richerson, P. J. and Boyd, R. (2005) *Not by genes alone: How culture transformed human evolution*, University of Chicago Press.

Robson, A. and Samuelson, L. (2011) The evolution of decision and experienced utilities, *Theoretical Economics*, **6**, 311–339.

Robson, A. J. (2001) Why would nature give individuals utility functions?, *Journal of Political Economy*, **109**, 900–914.

Robson, A. J. and Samuelson, L. (2007) The Evolution of Intertemporal Preferences, *The American Economic Review*, **97**, 496–500.

Robson, A. J. and Samuelson, L. (2010) The Evolutionary Foundations of Preferences, in *Handbook of Social Economics*.

Rünger, D. and Wood, W. (2015) Maintenance of Healthy Behaviors: Forming and Changing Habits, in *Behavioral Economics and Public Health*, Oxford University Press.

Samuelson, L. (2004) Information-Based Relative Consumption Effects, *Econometrica*, **72**, 93–118.

Samuelson, L. and Swinkels, J. M. (2006) Information, evolution and utility, *Theoretical Economics*, **1**, 119–142.

Sanfey, A. G. and Chang, L. J. (2008) Multiple Systems in Decision Making, *Annals of the New York Academy of Sciences*, **1128**, 53–62.

Sanfey, A. G., Loewenstein, G., McClure, S. M. and Cohen, J. D. (2006) Neuroeconomics: cross-currents in research on decision-making., *Trends in cognitive sciences*, **10**, 108–16.

Schelling, T. C. (1978) Egonomics , or the Art of Self-Management, *The American Economic Review*, **68**, 290–294.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J. and Griskevicius, V. (2007) The Constructive, Destructive, and Reconstructive Power of Social Norms, *Psychological Science*, **18**, 429–434.

Shu, L. L., Mazar, N., Gino, F., Ariely, D. and Bazerman, M. H. (2012) Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end, *Proceedings of the National Academy of Sciences*, **109**, 15197–15200.

Slemrod, J. (2007) Cheating ourselves: The economics of tax evasion, *Journal of Economic Perspectives*, **21**, 25–48.

Sozou, P. D. (1998) On hyperbolic discounting and uncertain hazard rates, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **265**, 2015–2020.

Spinella, M. (2003) Evolutionary mismatch, neural reward circuits, and pathological gambling, *International Journal of Neuroscience*, **113**, 503–512.

Stephens, D. W. (2002) Discrimination, discounting and impulsivity: A role for an informational constraint, *Philosophical Transactions of the Royal Society B: Biological Sciences*, **357**, 1527–1537.

Stephens, D. W. and Anderson, D. (2001) The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences, *Behavioral Ecology*, **12**, 330–339.

Tangney, J. P., Baumeister, R. F. and Boone, A. L. (2004) High self control predicts good adjustment, less pathology, better grades, and interpersonal success, *Journal of Personality*, **72**, 271–324.

Taubinsky, D. and Rees-Jones, A. (2018) Attention variation and welfare: theory and evidence from a tax salience experiment, *The Review of Economic Studies*, **85**, 2462–2496.

Teixeira, A. A. and Rocha, M. F. (2010) Cheating by economics and business undergraduate students: An exploratory international assessment, *Higher Education*, **59**, 663–701.

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

Tooby, J. and Cosmides, L. (1992) The psychological foundations of culture, in *The adapted mind: Evolutionary psychology and the generation of culture*, pp. 19–136.

Watson, K. K. and Platt, M. L. (2012) Social signals in primate orbitofrontal cortex, *Current Biology*, **22**, 2268–2273.

Williams, K. W. and Elmquist, J. K. (2012) From neuroanatomy to behavior: central integration of peripheral signals regulating feeding behavior, *Nature Neuroscience*, **15**, 1350–1355.

Zhu, L., Jenkins, A. C., Set, E., Scabini, D., Knight, R. T., Chiu, P. H. and King-casas, B. (2014) Damage to dorsolateral prefrontal cortex affects tradeoffs between honesty and self-interest, *Nature Publishing Group*, **17**, 1319–1321.

Appendix

Proof of Proposition 1. I will show that the following is an asymptotic equilibrium given beliefs $\beta(m)$: $v(y^2) = \frac{1}{\hat{f}(\beta^{-1}(y^2))}$, and $B = \mathbb{E}[\beta^{-1}(y^2)]$. System 2's problem:

$$\max_{(m_i)_{i=1}^K} \sum_{i=1}^{K} \hat{F}(m_i) \cdot v(y_i^2) \qquad s.t. \quad m_i \geq 0 \text{ for all } i, \text{ and } \quad \frac{1}{K}\sum_{i=1}^{K} m_i \leq B,$$

where $\hat{F}(m_i) = F(\beta(m_i))$. Recall that F is strictly concave, increasing and and differentiable, and so is β (by Assumption 3), and hence $\hat{F}$ is strictly concave and differentiable, and therefore the problem is well defined and the first order conditions for System 2 are sufficient for optimality. The first order conditions are then:

$$\hat{f}(m_i)v(y_i^2) = \lambda - \mu_i, \tag{3}$$

where μ_i corresponds to the Lagrange multiplier of the first constraint, and λ to that of the second constraint. The following equation solves System 2's condition for the asymptotic equilibrium:[27] $m_i = \hat{f}^{-1}\left(\frac{\lambda-\mu_i}{v(y_i^2)}\right)$. This guarantees that System 2 is maximizing its expected utility. In order to have an asymptotic equilibrium, it must be the case that the Principal's condition is also met, namely:

$$\lim_{K\to\infty} \mathbb{P}\left[(a_i)_{i=1}^K = \arg\max \sum_{i=1}^{K} y(a_i)\right] = 1.$$

But notice that the following events are all identical:

$$\left\{(a_i)_{i=1}^K = \arg\max \sum_{i=1}^{K} y(a_i)\right\} = \{a_i = 2 \Leftrightarrow y_i^2 > y_i^1 \ \forall i\} =$$

$$= \{\beta(m_i) > y_i^1 \Leftrightarrow y_i^2 > y_i^1 \ \forall i\} = \{\beta(m_i) = y_i^2 \ \forall i\} =$$

[27] Since $\hat{F}$ is increasing and concave, therefore $\hat{f}$ is decreasing, and hence $\hat{f}^{-1}$ is well defined.

$$\left\{\hat{f}^{-1}\left(\frac{\lambda-\mu_i}{v(y_i^2)}\right)=\beta^{-1}(y_i^2)\ \forall i\right\}=\left\{v(y_i^2)=\frac{\lambda-\mu_i}{\hat{f}(\beta^{-1}(y_i^2))}\ \forall i\right\}.$$

The last event can be further decomposed into three events

$$\left\{v(y_i^2)=\frac{\lambda-\mu_i}{\hat{f}(\beta^{-1}(y_i^2))}\ \forall i\right\}=\left\{v(y_i^2)=\frac{c}{\hat{f}(\beta^{-1}(y_i^2))}\ \forall i\right\}\cap\{c=\lambda\}\cap\{\mu_i=0\ \forall i\}, \tag{4}$$

for some constant $c>0$ (because λ is a positive Lagrange multiplier). In what follows, I take $c=1$, which is just a normalization. Hence, we must have $v(y_i^2)=\frac{1}{\hat{f}(\beta^{-1}(y_i^2))}$, and we need to prove that $\lim_{K\to\infty}\mathbb{P}[\{\lambda=1\}\cap\{\mu_i=0\ \forall i\}]=1$.

I will first show that the event $\{\lambda=1\}$ implies the event $\{\mu_i=0\ \forall i\}$. In order to do so, assume $\lambda=1$, and note that for a particular y_i^2 there are two options:

1. either the second constraint is not binding ($\mu_i=0$), in which case we have $\frac{\hat{f}(m_i)}{\hat{f}(\beta^{-1}(y_i^2))}=\lambda=1$,

2. or the second constraint is binding ($\mu_i>0$), in which case $m_i=0$ and we have that $\frac{\hat{f}(0)}{\hat{f}(\beta^{-1}(y_i^2))}=\lambda-\mu_i<1$, or equivalently $\hat{f}(0)<\hat{f}(\beta^{-1}(y_i^2))$. But since $\hat{f}$ is decreasing (because $\hat{F}$ is concave as discussed above), and $\beta^{-1}(y_i^2)\geq 0$ (since $y_i^2\geq 0$), this second case is not possible!

That is, if $\lambda=1$, the second constraint is never binding, and all $\mu_i=0$. Therefore, the event $\{\lambda=1\}$ implies the event $\{\mu_i=0\ \forall i\}$.

By the law of large numbers, we have that

$$\lim_{K\to\infty}\frac{1}{K}\sum_{i=1}^{K}m_i=\mathbb{E}[\hat{f}^{-1}((\lambda-\mu)\hat{f}(\beta^{-1}(y_i^2)))].$$

But because $\frac{1}{K}\sum_{i=1}^{K}m_i=B=\mathbb{E}[\beta^{-1}(y^2)]$, we have

$$\mathbb{E}[\hat{f}^{-1}((\lambda-\mu)\hat{f}(\beta^{-1}(y^2)))]=\mathbb{E}[\beta^{-1}(y^2)]. \tag{5}$$

But note that $\lambda = 1$ is a solution to Equation 5, because we already proved that $\lambda = 1$ implies $\mu_i = 0$ for all i, and thus $\lambda = 1$ solves $\mathbb{E}[\hat{f}^{-1}(\lambda \hat{f}(\beta^{-1}(y^2)))] = \mathbb{E}[\beta^{-1}(y^2)]$. And this concludes the proof that

$$\lim_{K\to\infty} \mathbb{P}\left[(a_i)_{i=1}^K = \arg\max \sum_{i=1}^K y(a_i)\right] = \lim_{K\to\infty} \mathbb{P}\left[\left\{v(y_i^2) = \frac{1}{\hat{f}(\beta^{-1}(y_i^2))}\ \forall i\right\} \cap \{\lambda = 1\} \cap \{\mu_i = 0\ \forall i\}\right] = 1,$$

and therefore of the existence of the asymptotic equilibrium. Notice that this equilibrium is essentially unique, in the sense that all the beliefs β produce the same outcome, namely System 1 chooses action 2 if and only if $\beta(m_i) \geq y_i^1$, or equivalently, if $y_i^2 = \beta(\beta^{-1}(y_i^2)) \geq y_i^1$. Thus, the beliefs do not affect the outcome of the equilibrium, and when we take $\beta(m) = m$, we obtain the forms of $v(y^2) = \frac{1}{f(y^2)}$ and $B = \mathbb{E}[y^2]$. Finally, note that no other λ fulfills the conditions in Equation 4, and thus there are no other asymptotic equilibria.

□

***Proof of Proposition** 2.* System 2's first order conditions are:[28]

$$f(m_1(\tau) - \rho(\tau))v(y_1^2) = \lambda(\tau),$$
$$f(m_j(\tau))v(m(y_j^2)) = \lambda(\tau) \qquad \text{for } j > 1.$$

Where $\lambda(\tau)$ is given as the solution to:

$$\rho(\tau) + \sum_{j=1}^K f^{-1}\left(\frac{\lambda(\tau)}{v(y_j^2)}\right) = B.$$

Because f^{-1} is decreasing in $\lambda(\tau)$, we have that $\lambda(S) > \lambda(C)$. Therefore, since $m_i(\tau) = f^{-1}\left(\frac{\lambda(\tau)}{v(y_i^2)}\right)$, we have that $m_i(C) > m_i(S)$ for all $i > 1$, and therefore System 1 chooses $a = 2$ more often in the Control, and so revealed self-control is larger under C than under S, as we

[28]For simplicity, I ignore the constraint $m_i \geq 0$ in this and the rest of proofs. As proved in the proof of Proposition 1, that constraint holds asymptotically with probability 1. I also assume $\beta(m) = m$.

wanted to show. □

Proof of Proposition 3. The distribution of $\bar{y}^2$ is given by a truncated normal, $\mathcal{N}_+\left(\theta, \frac{\sigma^2}{K}\right)$, with probability density function $\tilde{\phi}_+$. The normal distribution has the Monotone Likelihood Ratio Property (MLRP), which is inherited by the truncated normal:

$$\frac{d}{d\bar{y}^2}\left(\frac{\frac{d\tilde{\phi}_+}{d\theta}}{\tilde{\phi}_+}\right) > 0.$$

Let $p_t(\bar{y}^2(t)|\bar{y}^2(t-1))$ be the posterior of $\bar{y}^2(t)$, after having observed $y^2(t-1)$. Using Lemma 2 in Alonso *et al.* (2013), we have that $p_t(\bar{y}^2(t)|\bar{y}^2(t-1))$ inherits MLRP from $\tilde{\phi}_+$. Finally, note that

$$\frac{B(t)}{K} = \mathbb{E}[y^2] = \mathbb{E}[\bar{y}^2(t)] = \int \bar{y}^2(t)\ dp(\bar{y}^2(t)|\bar{y}^2(t-1)).$$

But MLRP implies First Order Stochastic Dominance, so $\int \bar{y}^2(t)\ dp(\bar{y}^2(t)|\bar{y}^2(t-1))$ is increasing in $\bar{y}^2(t-1)$, and therefore so is $B(t)$. □

Proof of Proposition 4. The proof is completely analogous to that of Proposition 1, except that the expectation of y^2 is conditional on γ when applying the law of large numbers, and thus the budget is $B(\gamma) = \mathbb{E}[y^2|\gamma]$.

□

Proof of Proposition 5. Let $W(B) = \mathbb{E}\left[\zeta \cdot u(a, y^2) + (1-\zeta) \cdot y^a\right]$, which is a function of the budget B (and implicitly of the actions a chosen a as a result of the behavior of Systems 1 & 2). Note that when the budget is $B = \mathbb{E}[y^2]$, then $\mathbb{E}[y^a]$ is maximized (by Proposition 1), and therefore $\frac{\partial \mathbb{E}[y^a]}{\partial B} = 0$. Note also that System 2's utility is increasing in a, and that a larger budget implies a higher frequency of choosing $a = 2$, and therefore $\frac{\partial \mathbb{E}[u(a,y^2)]}{\partial B} > 0$ for all B. Combining both comparative statics, it follows that $\frac{\partial W}{\partial B} > 0$ at $B = \mathbb{E}[y^2]$, and therefore the socially optimal budget B^* is larger than $\mathbb{E}[y^2]$. □

Proof of Proposition 6. If System 2 is naive, it still chooses m_i as if ρ was 0. Therefore, $\mathbb{P}[a = 2|\rho + y^1, y^2] = \mathbb{P}[\beta(m) > \rho + y^1] < \mathbb{P}[\beta(m) > y^1] = \mathbb{P}[a = 2|y^1, y^2]$.

If System 2 is sophisticated, it solves

$$\max \sum_{i=1}^{K} F(\hat{m}_i - \rho) \cdot v(y_i^2) \qquad s.t. \quad \frac{1}{K}\sum_{i=1}^{K} \hat{m}_i \leq B.$$

The optimality conditions are:

$$f(\hat{m}_i - \rho) \cdot v(y_i^2) = \hat{\lambda} \implies \hat{m}_i = \rho + f^{-1}\left(\frac{\hat{\lambda}}{v(y_i^2)}\right).$$

From the budget constraint, and the fact that $B = \mathbb{E}[y^2]$ and $v(y^2) = \frac{1}{f(y^2)}$, we know that

$$\frac{1}{K}\sum_{i=1}^{K} f^{-1}(\hat{\lambda} \cdot f(y_i^2)) = \mathbb{E}[y^2] - \rho.$$

Let $\tilde{m}_i$ and $\tilde{\lambda}$ be as in the original problem (i.e. when $\rho = 0$). Because $f^{-1}(\lambda \cdot f(y^2))$ is decreasing in λ, then we have that $\hat{\lambda} > \tilde{\lambda}$, and therefore that

$$\mathbb{P}[a = 2|\rho + y^1, y^2] = \mathbb{P}[\rho + f^{-1}(\hat{\lambda} \cdot f(y^2)) > \rho + y^1] < \mathbb{P}[f^{-1}(\tilde{\lambda} \cdot f(y^2)) > y^1] = \mathbb{P}[a = 2|y^1, y^2].$$

□

Proof of Proposition 7. Let $G_K(y^2)$ be the empirical distribution of y^2 when there are K choices. By the Glivenko-Cantelli theorem, we have that for all $\delta > 0$, there exists a $\bar{K}$ such that for all $K > \bar{K}$, $\sup_{y^2} |G_K(y^2) - G(y^2)| < \delta$ almost surely. Because System 2 chooses m_i as in Equation 2, that means that the average message sent by System 2 is bounded above by $(1+\delta)\mathbb{E}[y^2]$. Therefore, that means that there are a fraction of at most $1 - \frac{1}{1+\delta} = \frac{\delta}{1+\delta}$ tasks where its self-control is exhausted (and receives payoff of at least $u(1, y^2)$, since $v(y^2) = \frac{1}{f(y^2)} > 0$), so System 2's utility is bounded below by $\mathbb{E}[u(1, y^2)] + \frac{\mathbb{E}[\hat{F}(y^2) \cdot v(y^2)]}{1+\delta}$. Therefore given $\nu > 0$, we can find δ such that

$$\mathbb{E}[u(1,y^2)] + \frac{\mathbb{E}[\hat{F}(y^2) \cdot v(y^2)]}{1+\delta} = \mathbb{E}[u(1,y^2)] + \mathbb{E}[\hat{F}(y^2) \cdot v(y^2)] - \nu,$$

and solving for δ we find: $\delta = \frac{\nu}{\mathbb{E}[\hat{F}(y^2) \cdot v(y^2)] - \nu}$, where recall that $\mathbb{E}[u(1,y^2)] + \mathbb{E}[\hat{F}(y^2) \cdot v(y^2)]$ is System 2's expected utility in the static case. Therefore, choosing $\bar{K}$ such that for all $K > \bar{K}$, $\sup_{y^2} |G_K(y^2) - G(y^2)| < \delta$ almost surely, concludes the proof. □