The economics of innate rewards

Lucie Letrouit[1]

February 28, 2023

Abstract:

Neuroscientists and psychologists have brought to evidence the contribution of both *model-free* (or System I) and *model-based* (or System II) *valuation* mechanisms in the determination of individual decisions. While the former mechanisms typically unconsciously learn to associate actions with rewards (commonly dopamine rewards) based on past correlations, the latter typically rely on a cognitive model of the likely outcomes of actions and of their desirability. The balance between the two types of mechanisms is monitored by the brain depending on the internal and external environment of the decision (e.g. stress, complexity of the decision...). In this context, the contribution of this paper is fourfold: (1) I argue that the *decision utility* can be seen as an endogenously weighted average of an expected pleasure (or *innate reward*, which humans innately experience) and an expected *model-based utility* (which can rely on other fundamentals than pleasures, such as moral values). (2) I show that this framework allows to shed new light on a wide range of economic behaviors and to identify new sources of economic inefficiency in the typical intertemporal Social Welfare criteria, due to the nowadays large discrepancy between the *innate rewards* elicited by actions and these actions' desirability in the perspective of Humanity survival, as well as to the excessive role played by *innate rewards* in decision making in general. (3) I argue that the typical trade-offs implied by Welfare criteria, between present and future utility or present utility and *Humanity extinction risk*, could be strongly attenuated by the adoption of new types of policy measures aimed at (3.a) shaping *innate rewards* so that behaviors maximizing the *decision utility* also minimize the *Humanity extinction risk* and (3.b) reducing the role of *innate rewards* with respect to *model-based utility* in typical day-to-day economic decisions. (4) I briefly describe how such policy measures could be operationalized in practice.

Keywords: Decision making, Cognitive biases, Sustainable development, Extinction risk, Public policies

JEL classification: D91, D04, Q01, I31

[1]Gustave Eiffel University. Corresponding author: Lucie Letrouit, 14-20 Boulevard Newton Cité Descartes, Champs sur Marne F-77447 Marne la Vallée Cedex 2, lucie.letrouit@univ-eiffel.fr

1 Introduction

Neuroscientists and psychologists have brought to evidence the contribution of both *model-free* (also called System I as in Stanovich (1999)) and *model-based* (or System II) *valuation* mechanisms[2] in the determination of individual choices. In short, *model-free valuation* mechanisms typically unconsciously associate past actions with their rewards (commonly dopamine rewards), based on observed correlations, thereby leading individuals to repeat previously rewarded actions in a habitual manner. In contrast, *model-based valuation* mechanisms, which are prospective and goal-oriented, typically rely on a cognitive model of the likely outcomes of actions and of their desirability, given the knowledge acquired on the current environment, in order to determine the estimated overall desirability of an action.[3] It has been shown that the relative importance of these two kinds of mechanisms in decision making varies depending on the characteristics of the internal (i.e. internal state of the individual such as his level of hunger) and external (i.e. characteristics of the decision such as its required speed) environment in which the decision takes place. Additionally, recent evidence has shown that these *model- free* and *model-based valuation* mechanisms are not independent of each other. The *model-based valuation* mechanisms are able to modulate dopamine rewards so as to train *model-free valuations*.[4] The rewards perceived by the *model-free* mechanisms thus come both from *innate rewards*,[5] i.e. the naturally rewarding stimuli that need not be learned through experience (e.g. in response to fatty or sugary food, drugs, social interactions, novelty...), and *model-based induced rewards*, i.e. the modulations of rewards produced by the *model-based* mechanisms.[6]

Humans are endowed with *innate rewards* because those used to provide an evolutionary advantage to their hunter-gatherer ancestors, as they led the hunter-gatherers toward repeating actions that were good for their survival or their reproduction. For example, experiencing a reward when eating sugary food (such as an apple) or having sex led them to repeat this behavior in the future. Besides, the adaptability of the relative importance of *model-free* and *model-based valuation* mechanisms depending on the context was also evolutionarily advantageous as, the more depleted the needs of the hunter-gatherer were, the more he automatically relied on his *model-free valuations*, and thus on his *innate rewards*, which led him back to actions ensuring his survival and reproduction.[7] In a way, we may say that *innate rewards* played a role of "safety net", by ensuring that behaviors necessary for survival were performed. In contrast, nowadays, nature has largely been replaced by humans in the production of most consumption goods and their marketing, thereby leading to a decorrelation between the *innate rewards* (whose triggers have remained unchanged) and the actual *individual survival value* of goods.[8] For example, the excessive concentration of sugar in energy bars

[2] Some psychologists, like Evans and Stanovich (2013), call these mechanisms, respectively, Type I and Type II, because this denomination does not necessarily imply that "exactly two systems underlie the two forms of processing" (Evans and Stanovich, 2013).

[3] Evans and Stanovich (2013) argue that only two features are defining characteristics of the two types of processing, namely Type I processes are autonomous and require no working memory, while Type II processes are based on cognitive decoupling and mental simulation and require working memory. The other features generally associated with Type I and Type II processes are, they argue, "simply correlates that occur under well-defined conditions and are neither necessary nor defining features."

[4] See Moran et al. (2019); Doody, Van Swieten and Manohar (2022); Deserno et al. (2021); Gershman, Markman and Otto (2014); Doll et al. (2009); Doll, Simon and Daw (2012).

[5] Note that I will not, in this article, differentiate between rewards and pleasures, even though some literature suggests that dopamine rewards may not be necessary for "liking" and mostly implicated in the triggering of "motivation" for a specific behavior as well as learning (Berridge, 2009).

[6] Reciprocally, it can be thought that model-based valuations partially rely on innate rewards, as it is rational for individuals to value their pleasure, at the very least for their cognitive and emotional benefits, as positive affects have been shown to help in daily life for the planning and building of cognitive and emotional resources (Fredrickson et al., 2008; Dickinson et al., 2010).

[7] Imagine, for example, that the hunter-gatherer had been busy with painting the walls of a cave (thus relying on a *model-based valuation* of the informational or religious usefulness of his action), hunger would progressively increase his reliance on his *model-free valuation* mechanisms, which would, at some point, lead him to decide to go hunting or food gathering.

[8] The *individual survival value* of a an action (such as the consumption of a good) may be seen as a measure of the impact of this action on the individual's chances of survival.

is associated with a huge *innate reward*, under the form of a huge peak of dopamine, despite the unhealthiness of the repeated consumption of these bars. Similarly, pornographic videos' impact on the dopamine system has no common measure with its null or even negative impact on the individual's reproductive chances (Goetz et al., 2019), and the orchestrated excitement produced by marketing strategies around the goods they intend to sell (through music, discounts, loyalty programs, vivid colors...) often very efficiently dissimulates the total uselessness of the marketed products for the individual's survival and reproduction. At the opposite of the spectrum, the *innate costs* (i.e. negative *innate rewards*) faced by hunter-gatherers also used to sufficiently correctly reflect dangers for *individual survival.* For example, many dangerous food items have a bad taste. Nowadays, however, a large number of modern dangers are totally imperceptible to human senses so that they are associated with null *innate costs*, even if their *individual survival value* is strongly negative. We can think, for example, of a wide range of lethal chemicals, endocrine disruptors, radioactivity, asbestos, the increase of the CO_2 concentration in the atmosphere...[9] What's more, the *innate cost* linked with the act of purchase has been reduced to a ridiculously small level as consumers can now buy products online from home by a simple click and have them delivered right away without exerting any kind of effort. This development has metamorphosed *innate rewards* from natural allies for the survival of humans to natural deceivers, requiring from humans an ever increasing compensation of deceiving *innate rewards* by stronger reliance on *model-based valuations*, in order to limit or avoid unhealthy or survival-threatening behaviors.[10] Furthermore, Humanity has now reached a level of development that makes it possible for her to self-destroy if wrong decisions are taken. Inadequate *innate rewards* thus have acquired the potential to lead to Humanity's extinction. Indeed, the aggregation of overconsuming behaviors encouraged by excessive *innate rewards* associated with a wide range of products in every country of the world may very well lead to an irreversible climate change that will lead to Humanity's extinction (Pörtner et al., 2022).

In this context, the present paper's contribution is fourfold.

First, I propose to consider the *decision utility*[11] as a weighted average of a *model-free* and a *model-based utility* component, which weight varies depending on the internal and external environment of the decision. Then, focusing on the fundamentals leading to *model-free utility* and ignoring the errors linked with the computation of both utilities, I argue that we can further consider the *model-free utility* as a weighted average of two categories of rewards, namely *innate rewards* and *model-based induced rewards*, in accordance with the literature documented earlier in the introduction.[12][13]

Secondly, I show that this framework proves very useful to understand the role of *innate rewards* and *model-based valuations* in shaping a wide array of economic decisions, such as a typical purchase in a store or on a website or a typical choice between a "sustainable" and a "non-sustainable" action in daily life, depending on the individual's internal state. Furthermore, it allows to identify new sources of economic inefficiency in typical intertemporal Social Welfare criteria (such as the expected discounted utilitarian welfare criterion), due to the nowadays large discrepancy between the *innate rewards* elicited by actions and their desirability in the perspective of Humanity survival, as well as to the excessive role played by innate rewards in decision making.

Third, I argue, based on the proposed decomposition of the *decision utility*, that the typical

[9]Indeed, humans' senses and perceptions only allow them to apprehend a very tiny fraction of the characteristics of the world they live in and of the state of their own body. For example, they only perceive 0.0035% of the whole electromagnetic spectrum (Gawrylewski, Aug. 2019), they can't see things that are smaller than 0.1mm large, they perceive a flickering light of frequency higher than 60Hz as steady, and they lack many senses that other animals have (e.g. the perception of the Earth magnetic field, of electric fields...). A wide range

[10]See the literature on "evolutionary mismatches", including Li, van Vugt and Colarelli (2018) and Manus (2018).

[11]I will rely on the usual distinction in Behavioral Economics between *decision utility* (i.e. the neoclassical utility, namely the weight attributed to outcomes and attributes in decisions) and *experienced utility* (i.e. the pleasure and happiness really experienced as a consequence of the decisions).

[12]The fundamentals of *model-based utility* will not be explicitly modeled, but they typically include the *innate rewards* of the individual, the *innate rewards* of others, moral values, culture, scientific evidence...

[13]As errors in the computation of *model-based utility* are ignored, this utility is considered to perfectly reflect *expected utility.*

trade-offs implied by intertemporal Social Welfare criteria between present and future *experienced utility* or between present *experienced utility* and the *Humanity extinction risk* could be sizably attenuated by the adoption of one or more of the following new types of public policies: (1) policies improving the alignment of the pleasure-content (i.e. *innate reward*-content) of commercial products with their *Humanity survival value*,[14][15] (2) the limitation of factors leading to an excessive reliance on *innate rewards* (e.g. changing the shaping of purchase environments, reducing job-related stress to avoid stress traps...).

Fourth, I give some ideas on how such policy measures could be put in practice and, in particular, how *innate rewards* and *Humanity survival value* could be quantified for that purpose.

In the remaining of the paper, I first propose a brief literature review in Section 2. I then present the general analytical framework in Section 3, starting with the presentation of the model, followed by the identification of new sources of economic inefficiencies, the proposition of policy measures, and the discussion of their practical implementation. After that, I illustrate the general analytical framework with several applications of the model in Section 4 before concluding.

2 Literature

The present paper is linked with a wide array of literature strands from various research fields.

The literature that is the most akin to the present approach corresponds to a set of dual-process models of behavior proposed in the Behavioral Economics and Neuroeconomics fields. These models typically explain individual behaviors by the interaction between two different decision making systems or two different types of processing mechanisms, for example automatic versus control processes (Benhabib and Bisin, 2005) or planner-self versus doer-self (Thaler and Shefrin, 1981). They most often exclusively focus on intertemporal choice.[16] However, this is not the case of Loewenstein, O'Donoghue and Bhatia (2015) who propose a dual-process model based on a dichotomy between "deliberative processes" and "affective processes" and apply it not only to intertemporal choice but also to risky decision making and social preferences. While the distinction between *model-free* and *model-based valuation* mechanisms adopted in the present paper is reminiscent of previously made distinctions between System I and System II or automatic versus control cognitive processes, it departs from the literature by its focus on the fundamentals of valuations, namely *innate rewards* and *model-based valuations*, obtained through a further decomposition of model- free valuations into *innate rewards* and *model-based induced rewards*. The grounding of *model-free valuations* in *innate rewards* inherited from human evolution, which can be objectively quantified, then allows to discuss the economic inefficiencies associated with the pervasive discrepancy between *innate rewards* and *survival values* and to draw policy implications.

Relatedly, the present paper is also linked with the Psychology and Neuroscience literature dealing with the *model-free* (or System I or Type I) and *model-based* (or System II or Type II) *reinforcement learning* mechanisms that individuals rely on to learn the value of their actions. This literature attempts, in particular, to quantify the relative importance of *model-free* and *model-based valuations* in actual behaviors, depending on various internal and external environmental factors, based on laboratory experiments. Most papers largely rely on the methodology proposed by Daw et al. (2011). Among the numerous identified factors that affect the reliance of individual decisions on *model-free* mechanisms and *model-based* mechanisms, we can mention hunger (van Swieten, Bogacz

[14] The *Humanity survival value* of an action corresponds to its potential to increase Humanity's survival chances (i.e. reduce the *Humanity extinction risk*) or Humanity's future living conditions.

[15] Throughout the article, two functions are said to be "aligned" if they lead to exactly the same arbitrages between actions in all environments (if, according to one function, I prefer action A over action B, I will also prefer it according to the second function and vice versa, and if, according to one function, I choose a given quantity of an action, I will choose the same quantity according to the second function).

[16] See Thaler and Shefrin (1981); Shefrin and Thaler (1988); Bernheim and Rangel (2004); Fudenberg and Levine (2006). In this literature, one of the two "systems" is typically associated with a much larger discount rate than the other, which is used to explain the emergence and persistence of a number of addictions.

and Manohar, 2021), the complexity and uncertainty of the decision (Kim et al., 2019), as well as the stress level and size of working memory of the individual (Otto et al., 2013). The present paper relies on the insights of this empirical literature for the discussion of the model's implications.

Additionally, the present paper makes use of other insights from the neuroscience field regarding, in particular, the processing of *innate rewards* (Kelley and Berridge, 2002) and the existence of inter-individual differences in the intensity of these *innate rewards* (Reed and Knaapila, 2010).[17] It also relies on the Psychology and Medicine literature that discusses the development of "evolutionary mismatches", due to the growing discrepancy between *innate rewards* and the associated *individual survival values* (Li, van Vugt and Colarelli, 2018; Manus, 2018).

Furthermore, the economic literature that deals with the discussion of the adequate choice of long-term intertemporal Social Welfare function and the proposition of various optimality criteria[18] is mobilized, in order to identify the economic inefficiencies implied by the analytical framework proposed in this paper.

Eventually, the paper is related to the literature dealing with the translation of Behavioral economics' insights into public policy interventions. Behavioral economists, based on the observation of systematic gaps between *decision utility* and *experienced utility*, commonly argue that public policy should aim at maximizing the latter rather than the former. They argue that Behavioral Economics' results can make a significant contribution to the public debate "by offering new policy tools, improving predictions about the effects of existing policies, and generating new welfare implications" (Chetty, 2015). Among these new policy tools, they advocate for an improved shaping of default options, the use of anchors and reminders, as well as many other kinds of nudges and sludges (Thaler and Sunstein, 2009; Löfgren and Nordblom, 2020). In the present paper, the distinction between *decision utility* and *experienced utility* is mobilized. However, the gap between the two is solely explained by the intrusion of the *model-free valuation* (modulated by the level of need or stress) in decision making, while other contributors such as errors due to the use of heuristics are ignored.

3 Analytical framework

In this section, I start by presenting the general model, use it to derive new sources of economic inefficiency, propose policy measures aimed at addressing them and eventually describe how these policies could be implemented in practice.

3.1 Model

Consistently with the evidence presented in the Introduction and Literature Sections regarding the contribution of both *model-free* and *model-based valuation* in decision making, it will be assumed, in this model, that utility is a weighted average of a *model-free* and a *model-based valuation*.

Additionally, it has been explained that the pleasure rewards captured by the *model-free valuation* mechanisms may be of two types, the first of which corresponds to *innate rewards*, which humans have inherited from their evolution, due to the survival or reproduction advantage they gave to their owners. Humans thus experience *innate rewards* linked with sweetness and fatness because it helped their hunter-gatherer ancestors to recognize energy-rich food, sexual pleasure because it helped their ancestors want to reproduce... In this model, the notion of *innate rewards* will include *innate costs* (as negative *innate rewards*) corresponding, for example, to physical or intellectual efforts as well as pain, as they have been shown to also be encoded by dopamine neurons (Pasquereau and Turner, 2013; Varazzani et al., 2015; Ko and Wanat, 2016). Note that *innate rewards* are not necessarily self-centered. It has been shown that individuals have, to different extents, a capacity for empathy that leads them to actually "feel" the pleasures and displeasures of others, with actual hormonal signals similar to those that are released for self-centered pleasures and displeasures (Marsh, 2018).

[17] Note that this literature has shown that other neuromodulators are also implicated in reward processing, such as opioids and endocannabinoids (Berridge and Kringelbach, 2015), so that dopamine is not always necessary for *innate rewards* (Cannon and Bseikri, 2004).

[18] See Botzen and van den Bergh (2014); Méjean et al. (2020); Tsur and Zemel (2009); Stern (2006).

Further note that *innate rewards* appear to be modulated by many characteristics of the internal and external environment of the decision, such as the "homeostatic state, [...] the environment and epigenetics", as stated in Lewis et al. (2021) with numerous academic references.[19] For instance, humans experience more pleasure from eating if they are hungry.

The other category of pleasure rewards taken into account by *model-free valuation* mechanisms corresponds to *model-based induced rewards*. A number of publications[20] have indeed given evidence that *model-based* learning mechanisms can retrospectively update *model-free values* by modifying dopamine rewards (the *model-based induced rewards* then correspond to the modulations affecting dopamine rewards). Thus, in the proposed model, *model-free valuations* will correspond to a weighted average of *innate rewards* and *model-based induced rewards* (the latter encompassing, like the first one, its own negative version corresponding to *model-based induced costs*).

Model-based objectives (i.e. the fundamentals of *model-based valuations*) can be, for example, long-term healthiness, long-term popularity, the implementation of the individual's moral or religious values, the compliance with cultural or identity prescriptions, the improvement of the individual's offsprings' living conditions, the improvement of the individual's chances to meet his own basic needs in the future, becoming rich... But individuals may also choose as a *model-based* objective to maximize their (instantaneous or future) pleasure, in which case *model-free* and *model-based valuations* may become pretty similar. Note that although *model-based* objectives could be thought of as independent of the internal and external environment, the consequences of actions depend on the environment so that the *model-based valuation* of actions will depend on the environment too.

Eventually, the relative importance of the *model-free* and *model-based valuation* mechanisms in the final decision is arbitrated by neuronal computations (Lee, Shimojo and O'Doherty, 2014) and depends on numerous characteristics of the internal and external environment. For instance, it has been shown that the reliance on *model-free valuation* mechanisms increases with hunger (van Swieten, Bogacz and Manohar, 2021), the level of stress, especially for individuals with a low working memory (Otto et al., 2013), a reduced level of dopamine (Wunderlich, Smittenaar and Dolan, 2012), depression (Heo, Sung and Lee, 2021), drugs (Groman et al., 2019), the old age of the individual (Eppinger et al., 2013), the prosociality individuals (as opposed to selfishness) (Oguchi et al., 2023), the reduced processing speed of the brain (Schad et al., 2014), the combination of high complexity with high uncertainty of the decision (Kim et al., 2019), the required speed and accuracy of the decision (Keramati, Dezfouli and Piray, 2011), erotic cues (Mathar et al., 2022)... Some evidence suggests that this balance between *model-free* and *model-based valuation* mechanisms is not easy to train (Grosskurth et al., 2019).[21]

All in all, I propose to model the *decision utility* of an individual choosing an optimal combination of N actions $\boldsymbol{a} \in \mathbb{R}^{+N}$ in an (internal and external) environment characterized by the L-uple $\boldsymbol{s} \in \mathbb{R}^L$ as a weighted average of a *model-free utility* (or *model-free valuation*), $V^{mf}(\boldsymbol{a}, \boldsymbol{s})$, and a *model-based utility* (or *model-based valuation*), $V^{mb}(\boldsymbol{a}, \boldsymbol{s})$.[22] Both utilities will be assumed to be errorless, in the sense that the *model-free utility* will perfectly reflect the consequences of each combination of actions in terms of the (dopamine) rewards that will be activated[23] and the *model-based utility* will perfectly reflect the consequences of each combination of actions in terms of achievement of the individual's

[19] See also Watts and Bernat (2018).

[20] See Moran et al. (2019); Doody, Van Swieten and Manohar (2022); Deserno et al. (2021); Gershman, Markman and Otto (2014); Doll et al. (2009); Doll, Simon and Daw (2012).

[21] In figures, the relative importance of the *model-based valuation* estimated in Daw et al. (2011) is 39%, Wunderlich, Smittenaar and Dolan (2012) find 58% and Hackel et al. (2019) 83%. We can observe that it varies a lot depending on the experiment, be it due the characteristics of the population under study or the limited size of the samples.

[22] While this formulation of *decision utility* is clearly reminiscent of dual-self theories, the author is aware of the critiques formulated against them (Grayot, 2020) and rather seeks to translate the varying share of *model-free* and *model-based valuation* mechanisms implied in the decision making process. This approach does not require that the two types of mechanisms be associated with two clearly distinct systems in the brain nor exclude the possibility of their collaboration.

[23] This means that the fact that some environments or action choices have never been faced in the past and are thus associated with no definite *innate reward*, for example, is not taken into account.

model-based standards of what is desirable for him.[24][25] The *model-free* utility's weight, $\alpha(\boldsymbol{s}) \in [0,1]$, depends on the characteristics of the internal and external environment, $\boldsymbol{s}$. The *decision utility* of an individual thus writes:

$$U^{dec.}(\boldsymbol{s}) = \max_{\boldsymbol{a}} \alpha(\boldsymbol{s}) V^{mf}(\boldsymbol{a},\boldsymbol{s}) + (1-\alpha(\boldsymbol{s})) V^{mb}(\boldsymbol{a},\boldsymbol{s}) \tag{1}$$

The *model-free* utility can then be decomposed into the two sources of rewards it relies on, under the form of a weighted average of *innate rewards* $\mathcal{R}(\boldsymbol{a},\boldsymbol{s})$ and *model-based induced rewards*, with a weight $\beta \in [0,1]$. This weight will be considered to be constant across environments, as the author is aware of no article quantifying the magnitude of *model-based induced* modulations of rewards depending on the environment. Additionally, *model-based induced rewards* will be assumed to perfectly reflect *model-based valuations*, i.e. the training of the *model-free valuations* is errorless.[26] Consequently, the *decision utility* writes:[27][28]

$$U^{dec.}(\boldsymbol{s}) = \max_{\boldsymbol{a}} \alpha(\boldsymbol{s}) \beta \mathcal{R}(\boldsymbol{a},\boldsymbol{s}) + (1-\alpha(\boldsymbol{s})\beta) V^{mb}(\boldsymbol{a},\boldsymbol{s}) \tag{2}$$

Note that each individual will have his own $\mathcal{R}(.)$ and $V^{mb}(.)$ functions, depending on his specific sensitivities, genome, past experiences, moral values, knowledge (Lewis et al., 2021)... Each individual will also have his own $\alpha(.)$, depending, for example, on his genome (Mikus et al., 2022) and working memory size (Otto et al., 2013).[29] Inter-individual differences in $\alpha(.)$ have been shown to predict future behaviors, such as the evolution of their alcohol consumption three years later (Chen et al., 2021). Eventually, each individual will have his own β, translating the lower or higher ability of his *model-based* mechanisms to modify his *model-free valuations*.

This decomposition of the *decision utility* into *innate rewards* and *model-based valuations* voluntarily abstracts from the estimation errors linked with *model-free* and *model-based valuation* mechanisms, in order to focus on the fundamentals of these valuations, i.e. either "hard-wired" *innate rewards* that individuals are born with or individuals' *model-based* standards of what is desirable for them.[30] It is interesting because the two categories of fundamentals play very complementary roles for human survival. *Innate rewards* are based on very fine sensory signals captured by our

[24]This means that errors of reasoning, as well as lack of information or education leading to a bad evaluation of the consequences of a combination of actions will not be taken into account.

[25]Note that, in this model, the accuracy of *model-based valuations* is not explicitly featured and, a fortiori, not endogenous. Yet, it would seem natural that the more weight the brain decides to allocate to the *model-based valuation*, the more effort it will invest in order to make this valuation accurate. A number of papers seem to corroborate this intuition (Keramati, Dezfouli and Piray, 2011; Standage, Wang and Blohm, 2014). Additionally, the accuracy of both *model-free* and *model-based valuations* has been shown to depend on stress, which reduces the learning speed for both (Cremer et al., 2021).

[26]Note that, in reality, this may not be the case as there may be, for example, additional noise linked with the training mechanisms...

[27]Note that I assume, for the sake of simplicity, that all *model-based valuations* are used to train the *model-free* mechanisms, so that all combinations of actions and environment are associated with a *model-based induced reward*. Similarly, all combinations of actions and environments are associated with an *innate reward*, while, in reality, many of them can be expected to never have been experimented in the past and rather inferred from a gross association with a "similar" combination of actions in a "similar" environment.

[28]Also note that the interactions between the *model-free* and *model-based valuations* are most probably much more complicated than this. Indeed, the *model-based valuation* of outcomes relies, to various extents depending on the individual, on *innate rewards* (especially if the individual aspires for pleasures). It has already been stressed that it is rational for an individual to include *innate rewards* in his *model-based valuations* because of their positive impact on health, cognitive ability, etc. (Fredrickson et al., 2008; Dickinson et al., 2010). Additionally, *model-based* mechanisms can take into account, to some extent, the individual's *model-free* propensities (Moran, Keramati and Dolan, 2021). This could, theoretically, allow for the *model-based* mechanisms to anticipate on the *innate rewards* that will arise in a specific situation and calibrate themselves so as to offset them (in order to rule alone!). However, a full compensation would require that the individual fully measures, in each environment, the relative role played by his *model-free* mechanisms in his decisions, in order to be able to compensate for them in the future. In reality, individuals tend to rationalize their decisions, ex-post (Lind et al., 2017), so that *model-based* mechanisms can be expected to systematically downplay the role of *innate rewards*.

[29]Among rats also, differences in the propensity to approach signals versus goals have been documented (Flagel et al., 2007).

[30]As already noted, the *model-based* standards of an individual can be thought of as taking into account his own

organism, many of which we do not consciously analyze and thus do not integrate in the cognitive models on which our *model-based valuations* rely. For example, the facts that humans have a higher desire to procreate with sexual partners who have a complementary immune system (Kromer et al., 2016) or have a higher taste for protein-rich food when they are unwittingly made to lack proteins (Gibson, Wainwright and Booth, 1995) do not appear to be easily explained using solely *model-based valuations*. On the other side, *model-based valuations* make it possible for humans to anticipate mid-term and long-term consequences of their actions and to plan for the achievement of mid-term and long-term goals (such as working to obtain a diploma, saving to buy a house...), which would not be possible, based solely on *innate rewards*.

As already evoked in the introduction, the fact that the weight of the *model-free* mechanisms with respect to the *model-based* ones increases with the level of stress and need depletion also makes sense from an evolutionary point of view, as it ensured that hunter-gatherers came back to actions focused on the meeting of their most basic needs when necessary, by pursuing their *innate rewards*.[31]

Eventually, it can be remarked that the distinction between *model-free* and *model-based valuations* naturally provides two explanations at least for the gap between stated and revealed preferences (Haghani et al., 2021). First, when filling a form questioning them on their preferences, individuals will be far from all the little *innate rewards* scattered in their living environment and will not take them into account in their *model-free valuations*. Second, they will be less subject to need depletion or stress than in real life and will thus rely more (or even nearly exclusively) on their *model-based valuations*. This gap may be exploitable to quantify both *model-free* and *model-based valuations* in the real world, as will be discussed in Subsection 3.4.

As *model-based utility* is assumed to perfectly reflect the potential of a combination of actions to lead to the individual's desired outcomes, it can be interpreted as *expected utility*. *Innate rewards*, in contrast, are considered as short-term reward signals originally aimed at ensuring survival, but devoid of further value.[32] Thus:

$$U^{exp.}(\boldsymbol{s}) = \max_{\boldsymbol{a}} V^{mb}(\boldsymbol{a}, \boldsymbol{s}) \tag{3}$$

Remark: At this point, it can be interesting to momentarily assume (until the end of the Subsection only) that, in a given environment $\boldsymbol{s}$, each unit action is associated with a vector of elicited *innate rewards* (for example, for a unit action consisting in eating an ice cream in a restaurant with a group of friends for one unit of time, a first component of the vector would correspond to the pleasure linked with the sugary content of the food, a second component corresponding to the pleasure linked with the musical environment of the restaurant, a third to the pleasure linked with the positive social interactions with the friends sitting at the table...). We can additionally momentarily assume that a combination of actions simply leads to the same combination of the corresponding *innate rewards* vectors, and that all *innate rewards* dimensions are eventually aggregated (more precisely summed up) to produce the total *innate reward* function $\mathcal{R}(\boldsymbol{a}, \boldsymbol{s})$.[33] In mathematical terms, it would mean that:

innate rewards, but also, for example, the *innate rewards* of others (which he may have learned about by face-to-face interactions, or through books or through various media), some moral or religious or cultural values, and scientific evidence...

[31] As compared to Loewenstein, O'Donoghue and Bhatia (2015), "emotions" are not considered to make up the alternative mode of thinking to "rational reasoning" but rather to modulate the relative importance granted to the *model-free* and *model-based* mechanisms (and thus to *innate rewards* versus *model-based valuations*) through $\alpha(.)$, consistently with previously cited determinants of $\alpha(.)$ and with Huys and Renz (2017). Emotions are also allowed to potentially modulate *innate rewards* (through the $\boldsymbol{s}$ in $\mathcal{R}(\boldsymbol{a}, \boldsymbol{s})$) and *model-based valuations* (through the $\boldsymbol{s}$ in $V^{mb}(\boldsymbol{a}, \boldsymbol{s})$).

[32] But, as already noted, *model-based valuations* can grant some importance to pleasure experiences.

[33] Note that these simplifications clearly do not perfectly reflect reality, as the *innate rewards* associated with a unit action will clearly depend on the quantity of the action (a phenomenon of "satiation" or "boredom" will progressively take place). Besides, there may, in reality, be complementarities in the pleasurable content elicited by a combination of actions (two actions may, separately, yield less pleasure than when combined, for example cooking and eating as

$$\mathcal{R}(\boldsymbol{a},\boldsymbol{s}) = \mathbf{1}^t.\boldsymbol{R_s}.\boldsymbol{a} = (1,...,1).\begin{pmatrix} r^{\boldsymbol{s}}_{11} & r^{\boldsymbol{s}}_{12} & ... & r^{\boldsymbol{s}}_{1N} \\ r^{\boldsymbol{s}}_{21} & r^{\boldsymbol{s}}_{22} & ... & r^{\boldsymbol{s}}_{2N} \\ ... & ... & ... & ... \\ r^{\boldsymbol{s}}_{M1} & r^{\boldsymbol{s}}_{M2} & ... & r^{\boldsymbol{s}}_{MN} \end{pmatrix}.\begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_N \end{pmatrix} \tag{4}$$

where $\boldsymbol{a}$ is simply the vector of the quantity of each action in the combination of actions and $\boldsymbol{R_s}$ the matrix associating each action with its various *innate rewards*. For instance, $r^{\boldsymbol{s}}_{ij}$ is the *innate reward* of type i elicited by one unit of action j in the environment $\boldsymbol{s}$.

This formulation allows to reflect on the fact that, originally, in a hunter-gatherer society, the matrix $\boldsymbol{R_s}$ would be very sparse. Eating a fruit plucked from a tree only triggers few *innate rewards*, namely those that favor this evolutionarily advantageous action (such as the *innate reward* linked with the sugary content of the fruit). On the opposite, the matrix $\boldsymbol{R_s}$ in modern societies could be thought of as much denser and messy. A typical purchase situation in a store triggers a huge number of *innate rewards*, e.g. some linked with the musical environment, some linked with discounts, with bright colors, with the positive social interaction with the seller, with the sexual cues on the packaging of products... This means that, nowadays, most *innate rewards* triggering a purchase are totally out of their original evolutionary function. They add up and build *decision utilities* totally disconnected from the real *survival value* of the good.

3.2 New sources of economic inefficiency

In order to discuss the sources of economic inefficiency suggested by the model, I adopt a long-term intertemporal perspective. Many indicators of intertemporal Social Welfare have been developed by economists, in the course of the years, in particular in the objective to discuss optimal climate policies (See Botzen and van den Bergh (2014) for a review). Among them, the most well-known among economists corresponds to the expected discounted utilitarian welfare function, $W = \int_{t=1}^{+\infty} N(t)U(t)\rho^t dt$, where $N(t)$ is the total size of the population at t, $U(t)$ the expected utility of a representative individual and $\rho \in]0;1[$ the time discount factor. This indicator has been used, in particular, in many Integrated Assessment Models for climate policy, such as the DICE (Nordhaus, 1993) or the Stern (2006) Review. It is sometimes mobilized with alternative specifications including, for example, the Chichilnisky (2000) criterion,[34], the sustainable discounted utilitarianism criterion,[35] or the global welfare function.[36] Alternatives to the expected discounted utilitarian welfare that account for uncertainty include the Epstein-Zin utility[37] , the Maximin criterion,[38] and the Limited Degree of Confidence criterion.[39]

These indicators all include, more or less explicitly, either a trade-off between present and future utilities or a trade-off between utilities and the risk of Humanity extinction or both. In some approaches, the trade-off between utilities and the risk of Humanity extinction is even made very

compared to cooking and then eating the result) and there may also be complementarities in the aggregation of several dimensions of pleasurable content (the pleasure reward associated with a food item that is both fatty and sugary has been shown to be higher than the sum of the pleasure rewards associated with the same level of fat and sugar in two different food items, DiFeliceantonio et al. (2018)).

[34]It is a weighted average of a discounted sum of utilities, plus the "terminal utility value", which allows that neither the present nor the long-run future should dictate the criterion.

[35]It assigns a zero utility discount rate if and only if the present is better off than the future in an attempt to treat all generations alike (Dietz and Asheim, 2012).

[36]It assigns a larger weight to poor regions of the world, in an attempt to compensate for the tendency of Integrate Assessment Models to undervalue the welfare of these regions in global approaches (Tol, 2002).

[37]The Epstein-Zin utility (Epstein and Zin, 1991), which originated from the finance literature, allows for a separate calibration of risk aversion and time preferences:

$$U_t = [(1-\beta)c_t^{\rho} + \beta(E_t[U_{t+1}^{\alpha}])^{\rho/\alpha}]^{1/\rho} \tag{5}$$

where β is the utility discount factor and $1/(1-\rho)$ corresponds to the intertemporal elasticity of substitution.

[38]It focuses on the maximization of the welfare of the worst-off generation (Rawls, 2004).

[39]It is a combination of the expected utility theory and the maximin criterion (Chichilnisky, 2000).

explicit (Méjean et al., 2020; Tsur and Zemel, 2009).

Introducing in these Welfare criteria the distinction between *decision utility* and *experienced utility*, we could say that there is either a trade-off between present and future *experienced utilities* or a trade-off between *experienced utilities* and the risk of Humanity extinction or both. To the best of the author's knowledge, the main solutions proposed in the literature to attenuate these tradeoffs rely on technological progress or on a change in individuals' "values" (such as a reduction in consumerist values), which can be considered as *model-based valuations.*

In this paper, we will focus on two important sources of economic inefficiency, linked with the fact that both *decision utility* and *experienced utility* are outcomes of the economy. Economic inefficiency arises when (1) the economy produces *experienced utilities* that the produced *decision utilities* are not able to maximize, i.e. there is a non-alignment between *decision utilities* and *experienced utilities*,[40] or when (2) there is a discrepancy between the behaviors maximizing present *experienced utilities* and behaviors that would have maximized future *experienced utilities* and/or minimized the risk of Humanity extinction. This second discrepancy will be referred to as the "non-alignment between *experienced utilities* and *Humanity survival values*", where the *Humanity survival value* of an action corresponds to its potential to increase Humanity's survival chances (or Humanity's future living conditions).[41] Relying on the model presented previously, the first source of inefficiency can be decomposed into two sub-sources of inefficiency, namely the non-alignment of the *innate rewards* of actions with their *model-based valuations* and the excessive weight granted to *innate rewards* in decisions. In addition, the second source of inefficiency can be reinterpreted as a non-alignment between *model-based valuations* and *Humanity survival values.* These insights can be summarized in the following Proposition:

Proposition 3.2.1: *Recognizing that the shapes of decision utility and experienced utility functions are outcomes of the economy, two key sources of economic inefficiency are:*

1. *The non-alignment of decision utilities with experienced utilities, which can be decomposed into:*
   (a) *The non-alignment of the innate rewards of actions with model-based valuations.*
   (b) *The excessive weight of innate rewards in decision making.*
2. *The non-alignment between experienced utilities, i.e. model-based valuations, and Humanity survival values.*

There is clear evidence that these sources of economic inefficiency all coexist in present societies. Indeed, as previously documented in the paper, the *innate rewards* of actions, notably by their often excessive character, and the *innate costs* of actions, by their sometimes ridiculous smallness, have become widely disconnected from the *model-based valuations* that individuals may prefer to adopt. Additionally, the already cited literature on the determinants of the relative importance of *model-free* versus *model-based valuations* suggests many ways in which our economies may lead to an excessive reliance on *innate rewards*, for example due to excessive stress, fatigue or excitation. Eventually, it is clear that humans often adopt *model-based* goals that are not consistent with survival, be it their own and that of their offsprings or that of Humanity (we can think, for example, of suicide, and of individuals who embrace a career in a very polluting sport such as skydiving or parachuting).

[40]As already indicated, throughout the article, two functions are said to be "aligned" if they lead to exactly the same arbitrages between actions in all environments (if, according to one function, I prefer action A over action B, I will also prefer it according to the second function and vice versa, and if, according to one function, I choose a given quantity of an action, I will choose the same quantity according to the second function).

[41]It can be noted that the *Humanity survival value* of an action may differ from its *individual survival value*, i.e. the potential of an action to increase the survival and reproduction chances of the individual and his offsprings. Typically, actions that enhance the social status of an individual have a positive *individual survival value* but little *Humanity survival value.* On the contrary, actions that consist in renouncing to have offsprings, though devoid of *individual survival value* may have a *Humanity survival value*, due to the large CO_2 emissions linked with having a child (Murtaugh and Schlax, 2009).

To illustrate this Proposition on a specific example, the expected discounted utilitarian welfare criterion adopted in the Stern Review (Stern, 2006) can be slightly adapted. This criterion is expressed by $\int_{t=1}^{+\infty} N(t)U(t)e^{-\delta t}dt$, where $N(t)$ is the total population at time t, $U(t)$ is the (expected) utility of a representative individual at time t, and δ is the per unit of time risk that a catastrophe eliminates Humanity (or *Humanity extinction risk*) so that $e^{-\delta t}$ corresponds to Humanity's survival chances at the horizon of time t. It is maximized through an optimal choice of utility at each period of time. Starting from this criterion, the dependency of the utility and of the risk of extinction on the actions adopted by individuals at t is made explicit and the distinction between *decision utility* and *experienced utility* is introduced. Second, actions are written as the outcome of the *decision utility*'s maximization decision where the *decision utility*, $U_{\boldsymbol{m}}^{dec.}$, can be shaped by policy measures, $\boldsymbol{m} \in \mathbb{R}^M$, which are assumed to be kept constant through time. Third, the internal and external environment of the decision, $\boldsymbol{s}_t$, changes at each point of time. All in all, welfare can thus be written as:

$$W(\boldsymbol{m}) = \max_{\boldsymbol{m}} \int_{t=1}^{+\infty} N(t)U_{\boldsymbol{m}}^{exp.}(\boldsymbol{a_{mt}}, \boldsymbol{s}_t)e^{-\int_1^{+\infty}\delta(\boldsymbol{a_{m\tau}})d\tau}dt \text{ s.t. } \boldsymbol{a_{mt}} = \arg\max_{\boldsymbol{a}} U_{\boldsymbol{m}}^{dec.}(\boldsymbol{a}, \boldsymbol{s}_t) \tag{6}$$

Taking into the definition of the *decision* and *experienced utility* in the proposed model, we obtain:

$$\begin{aligned} W(\boldsymbol{m}) = \max_{\boldsymbol{m}} \int_{t=1}^{+\infty} & N(t)V_{\boldsymbol{m}}^{mb}(\boldsymbol{a_{mt}}, \boldsymbol{s}_t)e^{-\int_1^{+\infty}\delta(\boldsymbol{a_{m\tau}})d\tau}dt \\ & \text{s.t. } \boldsymbol{a_{mt}} = \arg\max_{\boldsymbol{a}} \alpha(\boldsymbol{s}_t)\beta\mathcal{R}_{\boldsymbol{m}}(\boldsymbol{a}, \boldsymbol{s}_t) + (1-\alpha(\boldsymbol{s}_t)\beta)V_{\boldsymbol{m}}^{mb}(\boldsymbol{a}, \boldsymbol{s}_t) \end{aligned} \tag{7}$$

Note that the *Humanity survival value* of an action in this framework is its ability to reduce the *Humanity extinction risk*. It thus corresponds, at t, to $e^{-\delta(\boldsymbol{a_{mt}})}$.

The three sources of economic efficiency identified in Proposition 3.2.1 translate here into: (1.a) $\boldsymbol{a_{mt}}$ does not maximize $V_{\boldsymbol{m}}^{mb}(\boldsymbol{a_{mt}}, \boldsymbol{s}_t)$, (1.b) $\alpha(\boldsymbol{s}_t)$ is too large, and (2) actions that maximize $V_{\boldsymbol{m}}^{mb}(\boldsymbol{a_{mt}}, \boldsymbol{s}_t)$ do not minimize $\delta(.)$ or, equivalently, do not maximize $e^{-\delta(.)}$.

3.3 Policy measures

In this article, *model-based valuations* (or, equivalently, *experienced utilities*) will be considered as fixed and already aligned with *Humanity survival values* through the implementation of usual economic policies aimed at the internalization of externalities such as taxes and subsidies...[42] The focus will be set on the discussion of additional policy measures aimed at altering the shaping and weighting of *innate rewards* so that the *decision utilities* align with *Humanity survival values.* The three sources of economic inefficiency listed in Proposition 3.2.1 now boil down to: the non-alignment of *innate rewards* with *Humanity survival values*[43] and the excessive weight given to *innate rewards*

[42]Note that, in reality, all *model-based* objectives are not necessarily conducive to a higher chance of survival or reproduction for the individual (let alone for Humanity), though they often include a dimension of transmission, albeit abstract, to other individuals, such as the next generations. Some individuals may indeed choose to sacrifice to prevent a bomb from killing many civilians, to fight against a dictator, to not get cured from a cancer in order to concentrate all their efforts on finishing a book... However, adequate taxes and subsidies can compensate for this and nevertheless align *model-based valuations* with *Humanity survival values.*

[43]As stressed in the introduction, *innate rewards* are clearly not aligned with *Humanity survival values.* Indeed, it is clear that, for our hunter-gatherer ancestors, *innate rewards* were largely in adequacy with *survival values* (more precisely, they were aligned with *individual survival values*, but, at that time, those did not differ much from *Humanity survival values* as humans had not reached the level of economic development necessary to be able to self-destroy), as they encouraged our ancestors to adopt survival-chances-enhancing behaviors. However, in present economies, *innate rewards* have become largely decorrelated from *survival values* (more precisely, they have become decorrelated from both *individual* and *Humanity survival values*), as humans have become able to produce a huge variety of ersatz associated with totally non-natural rewards (food can provide an agreeable sugary taste without any calories or,

as compared to *model-based valuations* in decision making. We can note that either aligning *innate rewards* with *Humanity survival values* or reducing to zero the weight of *innate rewards* in decisions suffices to reach optimality. However, none of the two can be perfectly realized in practice. That's why policy measures attending to both objectives are necessary. Some ideas regarding the form of these policies are presented hereafter.

First, policies aimed at aligning average *innate rewards* (computed over a representative sample of the population) with *Humanity survival values* will be able to take advantage of the fact that *innate rewards* tend to be normalized by the brain, in such a way that if the intensity of the maximum *innate rewards* is reduced, all *innate rewards* will be re-normalized upwardly.[44] These policies could take many forms. For example, in the domain of food, they could consist in regulations on the sugary and fatty content/density of aliments but also on the ersatz of sugar and fat that trigger similar rewards (e.g. low-calories sweeteners), as well as on all other ingredients that elicit excessive *innate rewards*. Typically, *innate rewards* should be aligned with the nutritional benefits of the aliments and the conversion coefficient between the two should be scaled so that no "excessive" *innate reward* is ever reached. Such a policy would be more efficient than a sin tax on the sugar or fat content, because a sin tax only appeals to the *model-based valuation* mechanisms, which keeps a wide avenue open for *innate rewards* to lead to excessive sugar and fat consumption. In the domain of entertainment, regulations on the *innate rewards* elicited by movies, TV series, video games... (e.g. *innate rewards* linked with violent content, sexual cues, speed of actions...) could also be implemented. As in the case of food, *innate rewards* could be aligned with the informative or educational value of the entertainment and scaled so that no excessive *innate reward* is ever reached. With regards to marketing strategies, regulations on the modalities under which firms can present their products for purchase to consumers could be introduced. For example, discounts, fidelity programs, musical environments in stores..., which all trigger *innate rewards*, could be banished as they occult the real *Humanity survival value* of the goods.[45] A number of typical interventions proposed by Behavioral Economists could also fall in this category of policy measures, such as default-choice design (insofar as it creates an *innate cost* of choosing the option with a lesser *Humanity survival value*), the use of graphic images and other prompts to influence purchasing decisions (as they can create *innate costs* for actions with a lower *Humanity survival value*, such as when images of unhealthy lungs are depicted on packs of cigarettes), or other nudges and sludges. They could indeed be used to counterbalance the excessive *innate rewards* associated with some goods using some additional *innate cost*, so as to align the net *innate reward* with the *Humanity survival value*.[46] Note that the calibration of *innate rewards* may also require the calibration of their time allocation, because time discounting is especially high for *innate rewards*. For this purpose, cognitive, as proposed by Lieder et al. (2019) could help encourage the persistent use of one "environmental-friendly" object over the purchase of a new "polluting" object (new clothes...). However, other methods could be devised. All in all, it must be noted that the alignment of *innate rewards* with *Humanity survival values* is, in some regard, a very natural preoccupation, as it is the logical *model-free* counterpart to advocating for taxes and subsidies (which affect *model-based valuations*) aimed at aligning individual utilities with social welfare, as is extensively done in the economic literature.

on the contrary, with too many of them, movies and series can procure *innate rewards* linked with social contact, without its benefits...). This leads humans to overconsume to a point that threatens the ability of their descendants to survive. Our present economies have thus transformed the "safety net" of *innate rewards* into what often resembles a "black hole" or a "trap", which, instead of inducing need-depleted, depressed or stressed individuals toward healthy behaviors, leads them toward goods associated with excessive, "unnatural" rewards, thereby impairing their ability to come back to a healthy state.

[44]This is suggested by a quite large body of evidence. See, notably, Schultz (2015) and Landry and Webb (2021).

[45]We can also remark that this kind of policy measures can be expected to have other positive impacts such as avoiding the dopamine desensitization that has been linked with the frequent consumption of excessive *innate rewards* (Davis et al., 2008; Nutt et al., 2015; Volkow et al., 2010; Imataka et al., 2022) and reducing the prevalence of addictions (Teegarden, Nestler and Bale, 2008; Carlin et al., 2016; Berridge and Robinson, 2016).

[46]Note that aligning *innate rewards* and *Humanity survival value* becomes even more important once it has been recognized that even our *model-based valuations* are largely built upon *innate rewards*.

Second, in cases where the average *innate reward* cannot be aligned with the *Humanity survival value*, it can make sense to implement policy measures aimed at reducing the relative weight of *innate rewards* as compared to *model-based valuations* in day-to-day decisions, as a second-best policy.[47][48] These policies could take the form of programs ensuring that the most basic needs of all inhabitants are provided for, so as to reduce stress and acute need depletion, thereby reducing reliance on *model-free valuations* both for the targeted individuals in the present and for their offsprings in the future.[49] They could also consist in normalizing purchasing decision environments (i.e. stores and commercial websites...) so that they favor reliance on *model-based valuations* rather than *innate rewards* (for example, purchasing decision environments should be relaxing, prevent hormonal imbalances and favor information over exciting cues for time-limited discounts...). Moreover, the preservation of the ability of individuals to rely on their *model-based valuation* processes also requires policy measures aimed at their protection from endocrine disruptors. Indeed, research has shown that exposure to lead or mercury during prenatal or early postnatal development can affect cognitive development in the long term, including the capacity to rely on *model-based valuation* mechanisms (Braun et al., 2011; Rauh et al., 2011; Gore et al., 2015).

Third, it may be possible to educate citizens with regards to their cognitive processes and their day-to-day manipulation, notably by firms and all types of institutions, and with regards to potentially efficient metacognition methods to compensate for it. However, current evidence seems to suggest that increasing *model-based* reliance through training is not evident (Grosskurth et al., 2019) and this kind of training would most probably only reach the most educated social classes. Thus, this category of policy can be expected to be much less efficient than the two first.

We can note that the policy measures presented in this Subsection may be interpreted in terms of "metapreferences" (Becker and Murphy, 1993). Indeed, all else being equal, individuals may generally be thought of as preferring to hold preferences that reduce the risk of human extinction. Thus, we can say that individuals hold "metapreferences" aligned with the minimization of the *Humanity extinction risk* δ and that policy measures aimed at aligning utilities with *Humanity survival value* simply aim at maximizing metautilities.

We can also remark that the objective to align utility with *survival values* does not have the same implications depending on the geographical scale under study. If, instead of studying the Humanity scale, which is the point of view adopted in this article, we focused on the scale of a specific country, the risks of extinction would not only be linked with global risks such as global warming, th occurrence of a worldwide epidemic, or a meteorite impact... they would also encompass, for example, political and geopolitical risks (including the risks to be invaded and wiped out by neighboring countries). Thus, the behaviors optimizing the *country survival value* would not always be aligned with those optimizing *Humanity survival values.* For instance, boosting consumption in a country, even if it is not necessary to increase utilities, may help national firms and thus increase the economic power of the country in the world economy, thereby increasing its geopolitical power. On the contrary, boosting consumption, if it is not necessary to increase utilities, does not make any sense at the scale of the whole Humanity and can be considered as having a negative *Humanity survival value* due to the environmental consequences. It implies that, ideally, the calibration of *innate rewards* should be decided for the whole world economy at once.

[47]It can even be seen as a first-best policy if *model-based valuations* are assumed to be perfectly aligned with *Humanity survival values.* Indeed, inter-individual differences in the perception of *innate rewards* will always prevent them to be perfectly aligned, for each individual, with *Humanity survival values*, so that reducing the relative weight of *innate rewards* remains desirable, even after having aligned average *innate rewards* with *Humanity survival values.*

[48]Note that increasing the weight granted to *model-based valuations* can additionally lead to a more precise *model-based valuation*, as suggested by a number of articles (Keramati, Dezfouli and Piray, 2011; Standage, Wang and Blohm, 2014), but this is not explicitly modeled here. Additionally, the accuracy of both *model-free* and *model-based valuations* has been shown to depend on stress, which reduces the learning speed for both (Cremer et al., 2021).

[49]Research has indeed shown that prenatal and early postnatal stress and nutrition have long-term effects on cognitive development, including reduced executive function and decision-making ability (Bock et al., 2015; Cohen et al., 2018).

3.4 In practice

So far, several concepts, such as the *Humanity survival value* of a good or the *innate rewards* associated with it have remained quite abstract. In this subsection, I attempt to describe how they could be estimated empirically.

We can note, first, that it is quite easy to detect the most excessive *innate rewards* in our economic environment, which are those that should be targeted by the proposed policy measures, as they tend to lead to the development of addictions in a portion of the population (e.g. drugs, sugary or fatty food, gambling, sex and pornography, internet, shopping, etc).[50] Then, to evaluate more precisely the *innate rewards* associated with a product, numerous methods used by neuroscientists can be mobilized. They allow to estimate, directly or indirectly, the dopamine release associated with the purchase of a product and its subsequent use. They include Positron Emission Tomography (PET) imaging, Functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG) and Magnetoencephalography (MEG), which are all already being used by neuromarketers to increase sales (Nilashi et al., 2020). Consequently, for each product, it could be theoretically possible to measure, on a representative sample of the population, the average *innate reward* elicited by its consumption. Yet, given the formidable number of products, some methodologies could be designed to reduce drastically the number of measures necessary. It could be possible, for example, to first quantify the characteristics of the product that are associated with *innate rewards* (for example its sugary and fatty content) and then to use conversion tables to translate these characteristics into the corresponding *innate rewards*. The conversion tables would have to be built, once and for all, in a laboratory setting and could then be applied for a large variety of products (they would associate, for example, various quantities of sugar and fat with their corresponding dopamine rewards).[51][52]

Then, as already stated, in particular in the Literature Section, methodologies have also been developed to evaluate the relative importance given to *model-free* and *model-based* control in decision making in a given environment (Deserno et al., 2015).

The estimation of the *Humanity survival value* of a good is more tricky as it involves the implicit comparison of a huge variety of goods of very different natures, both with respect to their function in the sustainment of human life and with respect to their Life-Cycle Cost, as well as the taking into account of Planetary Boundaries leading to difficult trade-offs in terms of environmental impacts. Yet, some indicators can nevertheless be imagined. For example, for food items, a measure of "nutritional indispensability" could be designed, indicating the quantity of vitamins, fibers and desirable nutrients... present per calory in the food, which could then be divided by the life-cycle cost of provisioning the product. The creation of such indicators would undoubtedly require much work and debate on the part of scientists, firms and legislators. That's why another, simpler and thus maybe more promising, approach could consist in aligning *innate rewards* with the average meta-preferences of individuals. These meta-preferences could indeed be computed using several distinct methods. First, surveys of "stated meta-preferences", whereby individuals would be asked to evaluate how much they think they should ideally value each category of goods with respect to each other, could be conducted.[53] Second, a method of "revealed meta-preferences" could be devised using the difference between the behaviors adopted by individuals when they go well (i.e.

[50]See Davis and Carter (2009); Gearhardt et al. (2011); Hartston (2012); Linnet et al. (2012); O'Sullivan et al. (2011); Voon et al. (2014).

[51]This would yield graphs associated with the consumption of various categories of goods as in Figure 3 of Panos and Baker (2010) for drugs.

[52]Note that the precise methodology to be adopted for these measures would have to be seriously debated among neuroscientists, as the dopamine responses differ a lot depending on the zone of the brain considered, the individual studied, the state of the individual studied, his age, the time scale adopted, the quantities of the goods that are being consumed...(Solinas et al., 2002; Marinelli and McCutcheon, 2014; Knutson et al., 2001).

[53]Besides, it would also be possible to directly ask individuals about their experiences of remorse after purchase or consumption of a good, which gives a good starting point to identify where gaps between *innate rewards* and *model-based valuations* lie.

when they have maximum *model-based* control) versus when they go bad (i.e. when they have maximum *model-free* control). In other words, "revealed meta-preferences" would be inferred from the difference between "revealed preferences" when individuals go well and "revealed preferences" when they go bad. None of these methodologies will be able to provide a perfect indicator of *Humanity survival value*. However, even imperfect indicators may be of great value in this case, if they lead to the moderation of the excessive (or "supranormal" in the terms of Goodwin, Browne and Rockloff (2015)) *innate rewards*.

To finish with, evaluations of the policy measures aimed at modifying *innate rewards* or changing the balance between *innate rewards* and *model-based valuations* could consist in measuring ex-post reported changes in declared happiness or in measuring declared happiness both before and after the implementation of the policy measures. The time delay between the implementation of the policy measure and its ex-post evaluation would however need to be long enough to allow for a certain adjustment of individuals' inner reward systems.

On the whole, the alignment of *innate rewards* with either proxied *Humanity survival values* or individuals' metapreferences does not appear to be such an unrealistic endeavor.

4 Applications

In this section, I present a few contexts in which the analytical framework proposed in this paper can bring interesting insights. The first application corresponds to a typical purchase environment in which a firm can choose both the *innate reward* associated with the purchasing of its product and the environment of the purchase, which influences the balance between *innate rewards* and *model-based valuations* in the decision making of its consumers. The second application is a typical daily choice of time allocation between two discrete actions associated with distinct *Humanity survival values* and inversely ranked *innate rewards*. The third one proposes an explanation of "poverty traps" as "stress traps", where a stress shock at one point in time both decreases long-term-oriented actions at that point (through an increased weight of *innate rewards* in decision making), but also increases stress in the future, thereby leading to a permanently higher reliance on *innate rewards* and a persistent and cumulative decrease in long-term-oriented actions. A few other applications are more briefly described at the end of the section.

4.1 Typical purchase environment

In this first example, let us consider a monopoly firm that sells one fixed type of good associated with an increasing and concave "consumption-related" *innate reward* function (perceived by the consumers when he consumes the good), $\mathcal{R}(x) \in \mathbb{R}$, where x is the quantity of the good consumed. The firm can choose the unit price of the good, $p \in \mathbb{R}^+$, the intensity of the net "purchase-related" *innate reward* perceived by a consumer when he decides to buy one additional unit of the good, γ,[54][55] which is assumed to be independent of the quantity purchased, and the environment in which purchases will be performed, parametrized by $s \in \mathbb{R}$ (the environment corresponds, typically, to the whole design of the store, or the whole design of the firm's web site).[56]

[54] This net *innate reward* aggregates the impact of discounts, fidelity cards, free shipping, music in the store, free shipping offers, minus the effort cost of queuing to pay...

[55] Tversky and Kahneman (1986) show that, even when we know that the outcome will be the same (for example if different discounts result in the same price reduction), we will still be more attracted to the higher discount percentage. Bertrand et al. (2010) find that showing the picture of an attractive woman increases demand for a loan by about as much as a 25% reduction in interest rate. Shaw and Bagozzi (2018) find that the endorsement of the product by a celebrity increases positive affect. And Sahni, Wheeler and Chintagunta (2018) show that adding the name of the message recipient to the email's subject-line increases the probability of the recipient opening it by 20%, which translates into an increase in sales leads by 31%.

[56] For example, vivid, highly saturated colors are perceived as exciting (Labrecque and Milne, 2012) and arousing (Gorn et al., 1997).

A representative consumer has an inner model of the net benefits he will enjoy if he buys a quantity x of the good, which translates into a *model-based valuation* $V^{mb}(x)$ that consists in the pure benefits of the purchase, $v(x)$ (an increasing and concave function of x), minus its cost px, so that $V^{mb}(x) = v(x) - px$. Note that, as stated in Subsection "Policy measures", the *model-based valuation* is assumed to be aligned with the *Humanity survival value*. Then, for the sake of simplicity, the "consumption-related" and "purchase-related" *innate rewards* are assumed to sum up, so that the *innate reward* perceived by the consumer when he purchases a quantity x of the good is $\mathcal{R}(x) + \gamma x$. The weight he unconsciously gives to his *innate rewards* as compared to his *model-based valuation* is directly drawn from the general analytical framework, with $\alpha(.) \in]0;1[$ an increasing function of s (s could be, for example, the level of consumer excitement triggered by the store). His *decision utility* can thus be written as:

$$U^{dec.}(s) = \max_{x \in \mathbb{R}^+} \alpha(s)\beta(\mathcal{R}(x) + \gamma x + \epsilon x) + (1 - \alpha(s)\beta)\left(v(x) - xp\right) \tag{8}$$

The first-order condition yields the quantity consumed: $\alpha(s)\beta(\mathcal{R}'(x)+\gamma+\epsilon)+(1-\alpha(s)\beta)\left(v'(x) - p\right) = 0$, from which the consumer's demand function, $D(p, \gamma, s)$, can be inferred.

The profit of the monopoly firm can be expressed as:

$$\Pi(p, \gamma, s) = pD(p, \gamma, s) - C(D(p, \gamma, s), \gamma, s) \tag{9}$$

where $C(x, \gamma, s)$ is the total cost incurred by the firm to produce a quantity x of the good.

Let us assume that $v(.)$, $\mathcal{R}(.)$ and $C(.)$ are such that Π is inverse-U shaped with respect to its three variables.[57] Then, we have the following Proposition:

Proposition 4.1.1:

- *For goods for which the innate reward is greater than the model-based valuation or smaller by a sufficiently small margin, the marketing strategy of the firm will attempt to draw the consumer toward his "model-free valuation" mechanisms with an "exciting" purchase environment and will mobilize a large quantity of purchase-related innate rewards (i.e. γ and s are large).*
- *For goods for which the innate reward is sufficiently lower than the model-based valuation, the marketing strategy of the firm will attempt to draw the consumer toward his "model-free valuation" mechanisms with a "relaxing" purchase environment and will mobilize fewer purchase-related innate rewards (i.e. γ and s are small).*

As relatively few categories of goods and services can be thought of as having a lower *innate reward* than *model-based valuation* in modern economies, this Proposition implies that excess purchase-related *innate rewards* perceived by consumers and permanent excess reliance of consumers on *innate rewards* as compared to *model-based valuations* will be a pervasive feature of purchase environments and lead to a generalized overconsumption. We all have in our minds modern stores full of stimuli, from music to discounts, bright colors, friendly faces on cereal boxes and all sorts of goods, (sometimes too) benevolent sellers... We can also think of the campaign organized by the firm Coca-Cola with soda distributors that distributed sodas for free to people who hugged them clearly attempted to draw consumers toward their *innate rewards*. Yet, other purchasing environments are much more relaxing. We can think about real estate agencies and banks for example. The model suggests that a reason for this is that the *model-based valuation* is, in this case, greater than *innate rewards*. Thus, it makes sense for the firm to encourage *model-based valuations*.

These two Propositions suggest that the manipulation of γ and s are complementary and that both should be prevented or compensated for.

[57] This makes sense, as a small investment aimed at increasing the *innate reward* associated with the purchase of a good or at shaping a *model-free*-reliance-increasing environment will increase demand and cost little, while large investments will, at some point, become powerless to push demand further up.

4.2 Typical daily choice of time allocation

Consider an individual who regularly faces a discrete choice between the two same actions, G and B (for "good" and "bad").[58] The bad action yields a higher *innate reward* than the good one, $\mathcal{R}_G < \mathcal{R}_B$, but the good action is associated with a higher *Humanity survival value* or, equivalently as they are aligned, a higher *model-based valuation* or *experienced utility*, $V_G^{mbH} > V_B^{mbH}$. Let us further assume that the environment in which the individual will have to take his day-to-day decision varies stochastically. More precisely, let us assume that the environment characteristic s follows a uniform distribution over the interval $[-\bar{S}, \bar{S}]$ (with $\alpha(-\bar{S}) = 0$ and $\alpha(\bar{S}) = 1$). Then, the individual will, when making the decision consciously, choose action B over action G if and only if:

$$\alpha(s)\beta\mathcal{R}_G + (1 - \alpha(s)\beta)V_G^{mbH} < \alpha(s)\beta\mathcal{R}_B + (1 - \alpha(s)\beta)V_B^{mbH} \tag{10}$$

i.e.

$$V_G^{mbH} - V_B^{mbH} < \frac{\alpha(s)\beta}{1 - \alpha(s)\beta}(\mathcal{R}_B - \mathcal{R}_G) \tag{11}$$

Proposition 4.2.1: *Provided that the gap between the innate rewards $\mathcal{R}_B$ and $\mathcal{R}_G$ is large enough with respect to the gap between the (perfect) model-based valuations V_G^{mbH} and V_B^{mbH}, more precisely provided that $\mathcal{R}_B - \mathcal{R}_G > \frac{1-\beta}{\beta}(V_G^{mbH} - V_B^{mbH})$, there will always be some environments s in which the individual will choose the "bad" action.*

In this context, a policy that reduces or reverses the ranking of the *innate rewards* associated with the two actions will increase the proportion of "good" behaviors and thus increase both Humanity survival chances and *experienced utility*.

At this point, we can note that the analytical framework proposed in this article only considers conscious decisions. Unconscious decisions may be thought of, caricaturally, as relying exclusively on the *model-free valuation* mechanisms (comprising both *innate rewards* and *model-based induced rewards*), without requiring the (conscious) *model-based valuation* mechanisms. This would be consistent with Gershman, Markman and Otto (2014)'s results, who observe behaviors that are "more consistent with a cooperative architecture in which the *model-free* system controls behavior, whereas the *model-based* system trains the *model-free* system by replaying and simulating experience", and with Kurdi, Gershman and Banaji (2019) who find that, while "explicit evaluations of novel targets are updated via *model-free* and *model-based* processes, implicit evaluations depend on the former but are impervious to the latter". Thus, the "unconscious decision utility" could be expressed as:

$$U^{dec.,uncons.}(s) = \max_{\boldsymbol{a}} V^{mf}(\boldsymbol{a}, s) = \max_{\boldsymbol{a}} \beta\mathcal{R}(\boldsymbol{a}, s) + (1 - \beta)V^{mbH}(\boldsymbol{a}, s) \tag{12}$$

where the *model-based* component only corresponds to *model-based induced rewards*, namely those created by the *model-based valuation* mechanisms to train the *model-free* mechanisms. In comparison with the (conscious) *decision utility*, only the relative weight of *innate rewards* with respect to *model-based valuations* is modified. The unconscious counterpart of Proposition 4.2.1 is thus:

Proposition 4.2.2: *Provided that the gap between the innate rewards $\mathcal{R}_B$ and $\mathcal{R}_G$ is large enough with respect to the gap between the (perfect) model-based valuations V_G^{mbH} and V_B^{mbH}, more precisely provided that $\mathcal{R}_B - \mathcal{R}_G > \frac{1-\beta}{\beta}(V_G^{mbH} - V_B^{mbH})$, the individual will always, when acting unconsciously, choose the "bad" action.*

[58] For example, the choice considered may be to bike or drive one's car to go to work, to mend one's clothes or buy new ones, to check in on our relatives or watch a TV series, to choose to cook a healthy meal or order a pizza online, to choose to take the train or the plane to go to a not-so-far-away city, to choose a close-by vacation place or to elect a far-away island where it is only possible to go by plane...

In this context, a policy that reduces or reverses the ranking of the *innate rewards* associated with the two actions will allow individuals to systematically unconsciously opt for "good" behaviors and, there also, increase both the Humanity survival chances and *experienced utility*.

For example, if we consider an individual faced with the decision of either mending their holed pullover or buying a new one, the *innate reward* linked with buying a new pullover, at least for individuals living in dense cities, may be seen as systematically exaggeratedly large (or the *innate cost* as exaggeratedly small) as compared to the *innate reward* (or the significant *innate cost*) linked with mending the pullover. In this respect, the possibility to purchase goods online and to have them delivered at home in no time has contributed to further reduce the *innate costs* associated with buying, as consumers do not have to walk or even drive to stores anymore to be able to make purchases.[59] On the opposite, the *Humanity survival value* linked with buying a new pullover can be seen as systematically smaller than the *Humanity survival value* linked with mending the pullover. Policies that increase the *innate cost* of purchasing online, for example simply by requiring from firms that, before completing a purchase, consumers solve some logic game or any kind of intellectual task could thus reduce the overconsumption tendency, especially if the *innate costs* is properly calibrated to compensate for the *innate rewards* assocated with the purchase (such as discounts, fidelity programs...).

4.3 Stress or need depletion traps

This third application shows how the model can shed new light on poverty traps. Let us consider an individual who, at each period of time t can choose between two possible actions: an instantaneously pleasurable action A (associated with the *innate reward* $\mathcal{R}_A > 0$ and a null *model-based valuation*) and a forward looking action F (associated with the *innate reward* $\mathcal{R}_F < 0$ and a longer run *model-based valuation* of $V_F^{mbH} > 0$ correctly aligned with its *Humanity survival value*). The level of risk for an individual to experience a shortage in the meeting of one of his basic needs (in which case he dies) is assumed to linearly decrease with the total time invested in the past in action F. For example, action F may correspond to "working" and action A to "alcohol drinking" or "TV watching". His level of stress at t is assumed to follow the same pattern and to be of the form: $s_t = \epsilon_t - s\int_0^t \mathbf{1}_{X_\tau=F} d\tau$ where X_t is the choice of action at time t and ϵ_t is drawn, at each t, from a uniform distribution on the interval $[\underline{S}, \bar{S}]$. The weight of *innate rewards* is assumed to increase with the level of stress s. As a consequence, the *decision utility* that governs the choice between action A and F at each t varies through time, depending on the level of stress:

$$U_t^{dec.} = \max_{X_t \in \{A,F\}} \alpha(s_t)\beta(\mathbf{1}_{X_t=A}\mathcal{R}_A + \mathbf{1}_{X_t=F}\mathcal{R}_F) + (1-\alpha(s_t)\beta)\mathbf{1}_{X_t=F}V_F^{mbH}$$

It is clear that we have the following Proposition:

Proposition 4.3.1: *A positive shock on the level of stress of the individual at a given period of time t_0 will not only directly increase the risk for the individual to die at each period, due to his increased chances of having chosen A over F at t_0, it will also indirectly increase this risk cumulatively in the future by permanently increasing future stress, and thus reliance on innate rewards and the probability to choose A over F at each future time t.*

This Proposition is corroborated by the observation that poor people have permanently higher levels of stress (higher cortisol levels) and focus more on immediate needs (Haushofer and Fehr, 2014; Shah, Mullainathan and Shafir, 2012). It is also in line with studies finding that poor people tend to adopt a number of rationally undesirable behaviors, using less preventive healthcare, failing to follow drug regimens, adopting more unhealthy behaviors, missing on appointments, working

[59] An aggravating factor linked with online purchases is that people will tend to fall back to the easy satisfaction of purchasing when their needs are particularly depleted or their stress particularly high, which will further increase their tendency to overrelie on their *innate rewards* and purchase more than necessary (Nederkoorn et al., 2009).

less productively, being less attentive parents, badly managing their finances,[60] as well as with the evidence that suggests that "minor situational details" play a large role for poor people in the final decision to recourse to banking and saving services or to enroll in social programs (Bertrand, Mullainathan and Shafir, 2004). As such, this mechanism can contribute to explain the long-term and disproportionately large impacts that negative shocks tend to have on the economic situation of the poor.[61]

In this context, both policy measures aimed at reducing stress, uncertainty and/or the drudgery of work (which leads to depleted needs) and policy measures aimed at aligning *innate rewards* with the actual *Humanity survival value* of actions are necessary to limit the long-term and self-reinforcing impacts of adverse shocks on the poorest. Note that policies deregulating labor markets or increasing competition within or between firms can be expected, on the contrary, to increase stress and thus to increase the focus of the poor on easy short-term solutions that bring high *innate rewards*, rather than on more long-term solutions that could help alleviate their poverty, which can turn out to be very detrimental in the longer run for the economy (Byron, Khazanchi and Nazarian, 2010).

Note that this mechanism may also contribute to explain the consistently worse economic outcomes of minority groups observed in many Western countries, insofar as these group experience higher stress due to their minority status (Sternthal, Slopen and Williams, 2011), which, in turn, may lead them to rely more on their *innate rewards* relative to their *model-based valuations*, as compared to members of the majority group, thereby leading them to under-invest in forward looking actions.

4.4 Other applications

The analytical framework can be useful in a number of other economic contexts, some of which I more briefly evoke hereafter.

First, the model highlights one of the costs associated with an excess diversity of choice within a given decision context. Indeed, while several very influential economic models rely more or less explicitly on the assumption that consumers always prefer more diversity of choice (e.g. Dixit and Stiglitz (1977)), psychologists have shown that there are several reasons why increasing choice diversity may reduce consumers' welfare. Chernev, Böckenholt and Goodman (2015), in their review of the literature, identify four factors that moderate the impact of choice diversity on choice overload, namely the "choice set complexity, decision task difficulty, preference uncertainty, and decision goal". Choice overload then translates into lower satisfaction/confidence of consumers with regard to their choice, as well as regret, choice deferral and higher switching likelihood. The present model suggests that a high complexity and uncertainty of choice can lead consumers to fall back, momentarily, on their *innate rewards*, which leads them to regret their decision ex-post, when the *model-based valuations* regain more weight in their decision making. The model thus suggests that purchase environments, in order to favor *model-based valuations*, should also be designed in a way that moderates the complexity of each decision or so as to helps consumers deal with this complexity, thereby preventing them to fall back on their *innate rewards*. Public policies could be designed to try to meet this objective.

Second, the model can be linked with the phenomena of fashion and fads. Indeed, humans have evolved to be curious of their environment and to experience a dopamine release in association with novel or unexpected stimuli (Costa et al., 2014). Thus, their *innate rewards* will systematically lead them to overvalue anything new or that has not been seen in a long time. This can contribute to explain cyclical fashion, that arises even in the absence of any increase in the quality of products (and

[60] See Katz and Hofer (1994); DiMatteo et al. (2002); Cutler and Lleras-Muney (2010); Pampel, Krueger and Denney (2010); Karter et al. (2004); Neal et al. (2001); Kim, Sorhaindo and Garman (2006); McLoyd (1998); Barr (2012); Blank and Barr (2009); Edin and Lein (1997).

[61] See Hallegatte et al. (2020)

thus of any increase of their *model-based value*),[62] and certainly constitutes a barrier to sustainable development.

More generally, the model has implications for the analysis of innovations. Indeed, innovations can affect both the *innate rewards* and the *model-based valuations* of products. An innovation that affects the *innate rewards* of a product can, for instance, be an innovation on the design of the product (e.g. colors, global shape, location of buttons...), which will make it more appealing to the consumer's senses. In contrast, an innovation that affects the *model-based valuation* of a product is, for example, an innovation by which the functionalities of the product are improved (e.g. computer with a higher Random Access Memory (RAM) or a better graphics card). An innovation that only improves the *model-based valuation* of a product may not have such a large success, if it is not complemented by an *innate rewards* innovation (like a new design for the higher technology computer). But an *innate rewards* innovation can by itself generate a new fashion or trend (as can be observed from the fast fashion phenomenon for clothes). Innovations could thus be classified depending on whether they affect more the *innate rewards* or *model-based valuations* of consumers. And, when seeking to align *innate rewards* with the *Humanity survival value*, the *innate rewards* linked with novelty should be taken into account. For example, the *innate rewards* due to novelty could be regulated to reflect the real *Humanity survival value* of the embedded innovation.

Eventually, the model's implications with regards to the internalization of externalities are also of some interest. Numerous articles from the functional neuroimaging literature have provided strong evidence that humans rely on empathic simulations in order to represent in their brain the internal states of others, so that they actually "feel" for real the pain of others.[63] Thus, there are *innate costs* to hurting others' interests. However, the actual magnitude of these costs with respect to the real costs imposed on others may vary widely, depending both on the situation (e.g. the perceived proximity to those that may be hurt by one's actions...) and on the characteristics of the individual making the decision (e.g. level of empathy...). For example, it has been shown that humans tend to exhibit more empathy toward people they perceive as in-group members as compared to out-group members (Cikara, Bruneau and Saxe, 2011) and that they differ a lot with regards to their levels of empathy towards others in general (Jolliffe and Farrington, 2006). Individuals may also hold very different *model-based valuations* of their impact on others' welfare, depending on their moral values, culture, religion... (Schwartz, 2006). Although inter-individual differences may of course prove very complex to take into account in the implementation of policy measures, it could nevertheless be possible, theoretically, to calibrate *innate rewards* in a way that takes into account an already prevailing average level of empathy in the economic situation at stake and to modulate taxes aimed at the internalization of externalities (i.e. *model-based costs*) depending on the average level of altruistic norm in this very economic situation.[64]

5 Conclusion

The present paper invites to reflect on the huge effect that *innate rewards* have on human decisions, and to contemplate the fact that, instead of being harnessed by policy makers in the objective to improve Humanity's chances of survival, they are disproportionately appropriated by firms in the objective to intensify consumption behaviors at the expense of other (non-commercial) behaviors such as discussions, out-door playing... The paper underlines that, whereas *innate rewards* used to play the role of a "safety net" for our hunter-gatherer ancestors, in the sense that, in situations

[62] Other theoretical explanations of fashion cycles include a logic of social differentiation (Blumer, 1969; Berger and Heath, 2007).

[63] Several zones of the brain, collectively called the "pain matrix", are activated both by firsthand pain and the observation of the pain of others (Marsh, 2018).

[64] For example, taxes, according to this logic, should depend on the country, as it has been shown that the levels of altruism differ across countries (Schwartz, 2006).

of stress or need depletion, their increased role led hunter-gatherers to concentrate on their basic needs by adopting survival-chances-enhancing actions (like searching for health, natural food, or for a sexual partner), *innate rewards* nowadays rather play the role of a "black hole", whereby stressed or need-depleted individuals are led by excessive *innate rewards* or too weak *innate costs* toward survival-chances-damaging actions (like eating over-sugary or over-fatty food, binge watching TV series, or manipulating unwittingly objects made of dangerous chemical components). It suggests that governments' use of unconscious decision making mechanisms could be much more extensive than the mere implementation of nudges to increase the use of public services and the compliance with regulations. It could shape, for example, the interfaces between firms and customers, so as to limit the *innate rewards* elicited by acts of purchase, and it could shape the *innate rewards* elicited by the use of purchased products (for example through regulation on the sugary content of food). These actions would allow to realign *innate rewards* with the *Humanity survival value* of the products. Policy measures could also try, as a second-best, to encourage reliance on *model-based valuations*, rather than *innate rewards*, in day-to-day decisions, for instance by decreasing stress and uncertainty.

While envisioning human decision making as purely or mostly rational with only minimal cognitive biases may lead some economists to consider these kinds of policy interventions as paternalistic, the author thinks that the *model-free / model-based valuation* distinction and its grounding in evolutionary theory helps understand the key role played by *innate rewards* in our daily decisions, as well as their always increasing manipulation by firms. In this context, the author believes that policy measures should address the divide between *innate rewards* and the actual *survival value* of goods and attempt to reduce it, as it could powerfully bend human behaviors toward more sustainable-development-compatible activities and thereby non-negligibly contribute to the mitigation of global warming.

6 Bibliography

Barr, Michael S. 2012. *No slack: The financial lives of low-income Americans.* Brookings Institution Press.

Becker, Gary S, and Kevin M Murphy. 1993. "A simple theory of advertising as a good or bad." *The Quarterly Journal of Economics*, 108(4): 941–964.

Benhabib, Jess, and Alberto Bisin. 2005. "Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption–saving decisions." *Games and economic Behavior*, 52(2): 460–492.

Berger, Jonah, and Chip Heath. 2007. "Where consumers diverge from others: Identity signaling and product domains." *Journal of consumer research*, 34(2): 121–134.

Bernheim, B Douglas, and Antonio Rangel. 2004. "Addiction and cue-triggered decision processes." *American economic review*, 94(5): 1558–1590.

Berridge, Kent C. 2009. "'Liking'and 'wanting'food rewards: brain substrates and roles in eating disorders." *Physiology & behavior*, 97(5): 537–550.

Berridge, Kent C, and Morten L Kringelbach. 2015. "Pleasure systems in the brain." *Neuron*, 86(3): 646–664.

Berridge, Kent C, and Terry E Robinson. 2016. "Liking, wanting, and the incentive-sensitization theory of addiction." *American Psychologist*, 71(8): 670.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. 2010. "What's advertising content worth? Evidence from a consumer credit marketing field experiment." *The quarterly journal of economics*, 125(1): 263–306.

Bertrand, Marianne, Sendhil Mullainathan, and Eldar Shafir. 2004. "A behavioral-economics view of poverty." *American Economic Review*, 94(2): 419–423.

Blank, Rebecca M, and Michael S Barr. 2009. *Insufficient funds: Savings, assets, credit, and banking among low-income households.* Russell Sage Foundation.

Blumer, Herbert. 1969. "Fashion: From class differentiation to collective selection." *The sociological quarterly*, 10(3): 275–291.

Bock, Joerg, Tamar Wainstock, Katharina Braun, and Menahem Segal. 2015. "Stress in utero: prenatal programming of brain plasticity and cognition." *Biological psychiatry*, 78(5): 315–326.

Botzen, WJ Wouter, and Jeroen CJM van den Bergh. 2014. "Specifications of social welfare in economic studies of climate policy: overview of criteria and related policy insights." *Environmental and Resource Economics*, 58: 1–33.

Braun, Joe M, Amy E Kalkbrenner, Antonia M Calafat, Kimberly Yolton, Xiaoyun Ye, Kim N Dietrich, and Bruce P Lanphear. 2011. "Impact of early-life bisphenol A exposure on behavior and executive function in children." *Pediatrics*, 128(5): 873–882.

Byron, Kristin, Shalini Khazanchi, and Deborah Nazarian. 2010. "The relationship between stressors and creativity: a meta-analysis examining competing theoretical models." *Journal of Applied Psychology*, 95(1): 201.

Cannon, Claire Matson, and Mustafa R Bseikri. 2004. "Is dopamine required for natural reward?" *Physiology & behavior*, 81(5): 741–748.

Carlin, Jesse L, Sarah E McKee, Tiffany Hill-Smith, Nicola M Grissom, Robert George, Irwin Lucki, and Teresa M Reyes. 2016. "Removal of high-fat diet after chronic exposure drives binge behavior and dopaminergic dysregulation in female mice." *Neuroscience*, 326: 170–179.

Chen, Hao, Negin Mojtahedzadeh, Matthew J Belanger, Stephan Nebe, Sören Kuitunen-Paul, Miriam Sebold, Maria Garbusow, Quentin JM Huys, Andreas Heinz, Michael A Rapp, et al. 2021. "Model-based and model-free control predicts alcohol consumption developmental trajectory in young adults: A 3-year prospective study." *Biological psychiatry*, 89(10): 980–989.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman. 2015. "Choice overload: A conceptual review and meta-analysis." *Journal of Consumer Psychology*, 25(2): 333–358.

Chetty, Raj. 2015. "Behavioral economics and public policy: A pragmatic perspective." *American Economic Review*, 105(5): 1–33.

Chichilnisky, Graciela. 2000. "An axiomatic approach to choice under uncertainty with catastrophic risks." *Resource and Energy Economics*, 22(3): 221–231.

Cikara, Mina, Emile G Bruneau, and Rebecca R Saxe. 2011. "Us and them: Intergroup failures of empathy." *Current Directions in Psychological Science*, 20(3): 149–153.

Cohen, Juliana FW, Sheryl L Rifas-Shiman, Jessica Young, and Emily Oken. 2018. "Associations of prenatal and child sugar intake with child cognition." *American journal of preventive medicine*, 54(6): 727–735.

Costa, Vincent D, Valery L Tran, Janita Turchi, and Bruno B Averbeck. 2014. "Dopamine modulates novelty seeking behavior during decision making." *Behavioral neuroscience*, 128(5): 556.

Cremer, Anna, Felix Kalbe, Jan Gläscher, and Lars Schwabe. 2021. "Stress reduces both model-based and model-free neural computations during flexible learning." *NeuroImage*, 229: 117747.
Cutler, David M, and Adriana Lleras-Muney. 2010. "Understanding differences in health behaviors by education." *Journal of health economics*, 29(1): 1–28.
Davis, Caroline, and Jacqueline C Carter. 2009. "Compulsive overeating as an addiction disorder. A review of theory and evidence." *Appetite*, 53(1): 1–8.
Davis, Jon F, Andrea L Tracy, Jennifer D Schurdak, Matthias H Tschöp, Jack W Lipton, Deborah J Clegg, and Stephen C Benoit. 2008. "Exposure to elevated levels of dietary fat attenuates psychostimulant reward and mesolimbic dopamine turnover in the rat." *Behavioral neuroscience*, 122(6): 1257.
Daw, Nathaniel D, Samuel J Gershman, Ben Seymour, Peter Dayan, and Raymond J Dolan. 2011. "Model-based influences on humans' choices and striatal prediction errors." *Neuron*, 69(6): 1204–1215.
Deserno, Lorenz, Quentin JM Huys, Rebecca Boehme, Ralph Buchert, Hans-Jochen Heinze, Anthony A Grace, Raymond J Dolan, Andreas Heinz, and Florian Schlagenhauf. 2015. "Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making." *Proceedings of the National Academy of Sciences*, 112(5): 1595–1600.
Deserno, Lorenz, Rani Moran, Jochen Michely, Ying Lee, Peter Dayan, and Raymond J Dolan. 2021. "Dopamine enhances model-free credit assignment through boosting of retrospective model-based inference." *Elife*, 10: e67778.
Dickinson, Anthony, Bernard Balleine, M Kringelbach, and K Berridge. 2010. "Pleasures of the Brain."
Dietz, Simon, and Geir B Asheim. 2012. "Climate policy under sustainable discounted utilitarianism." *Journal of environmental economics and management*, 63(3): 321–335.
DiFeliceantonio, Alexandra G, Géraldine Coppin, Lionel Rigoux, Sharmili Edwin Thanarajah, Alain Dagher, Marc Tittgemeyer, and Dana M Small. 2018. "Supra-additive effects of combining fat and carbohydrate on food reward." *Cell metabolism*, 28(1): 33–44.
DiMatteo, M Robin, Patrick J Giordani, Heidi S Lepper, and Thomas W Croghan. 2002. "Patient adherence and medical treatment outcomes a meta-analysis." *Medical care*, 794–811.
Dixit, Avinash K, and Joseph E Stiglitz. 1977. "Monopolistic competition and optimum product diversity." *The American economic review*, 67(3): 297–308.
Doll, Bradley B, Dylan A Simon, and Nathaniel D Daw. 2012. "The ubiquity of model-based reinforcement learning." *Current opinion in neurobiology*, 22(6): 1075–1081.
Doll, Bradley B, W Jake Jacobs, Alan G Sanfey, and Michael J Frank. 2009. "Instructional control of reinforcement learning: a behavioral and neurocomputational investigation." *Brain research*, 1299: 74–94.
Doody, Max, Maaike MH Van Swieten, and Sanjay G Manohar. 2022. "Model-based learning retrospectively updates model-free values." *Scientific Reports*, 12(1): 2358.
Edin, Kathryn, and Laura Lein. 1997. *Making ends meet: How single mothers survive welfare and low-wage work.* Russell Sage Foundation.
Eppinger, Ben, Maik Walter, Hauke R Heekeren, and Shu-Chen Li. 2013. "Of goals and habits: age-related and individual differences in goal-directed decision-making." *Frontiers in neuroscience*, 7: 253.
Epstein, Larry G, and Stanley E Zin. 1991. "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis." *Journal of political Economy*, 99(2): 263–286.
Evans, Jonathan St BT, and Keith E Stanovich. 2013. "Dual-process theories of higher cognition: Advancing the debate." *Perspectives on psychological science*, 8(3): 223–241.
Flagel, Shelly B, Stanley J Watson, Terry E Robinson, and Huda Akil. 2007. "Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats." *Psychopharmacology*, 191: 599–607.
Fredrickson, Barbara L, Michael A Cohn, Kimberly A Coffey, Jolynn Pek, and Sandra M Finkel. 2008. "Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources." *Journal of personality and social psychology*, 95(5): 1045.
Fudenberg, Drew, and David K Levine. 2006. "A dual-self model of impulse control." *American economic review*, 96(5): 1449–1476.
Gawrylewski, Andrea. Aug. 2019. "A Whole New View." *Scientific American Space & Physics 2, 4, 2.*
Gearhardt, Ashley N, Sonja Yokum, Patrick T Orr, Eric Stice, William R Corbin, and Kelly D Brownell. 2011. "Neural correlates of food addiction." *Archives of general psychiatry*, 68(8): 808–816.
Gershman, Samuel J, Arthur B Markman, and A Ross Otto. 2014. "Retrospective revaluation in sequential decision making: a tale of two systems." *Journal of Experimental Psychology: General*, 143(1): 182.

Gibson, EL, CJ Wainwright, and DA Booth. 1995. "Disguised protein in lunch after low-protein breakfast conditions food-flavor preferences dependent on recent lack of protein intake." *Physiology & behavior*, 58(2): 363–371.

Goodwin, BC, Matthew Browne, and Matthew Rockloff. 2015. "Measuring preference for supernormal over natural rewards: a two-dimensional anticipatory pleasure scale." *Evolutionary Psychology*, 13(4): 1474704915613914.

Gore, Andrea C, Vesna A Chappell, Suzanne E Fenton, Jodi Anne Flaws, Angel Nadal, Gail S Prins, Jorma Toppari, and R Thomas Zoeller. 2015. "EDC-2: the Endocrine Society's second scientific statement on endocrine-disrupting chemicals." *Endocrine reviews*, 36(6): E1–E150.

Gorn, Gerald J, Amitava Chattopadhyay, Tracey Yi, and Darren W Dahl. 1997. "Effects of color as an executional cue in advertising: They're in the shade." *Management science*, 43(10): 1387–1400.

Grayot, James D. 2020. "Dual process theories in behavioral economics and neuroeconomics: A critical review." *Review of Philosophy and Psychology*, 11(1): 105–136.

Groman, Stephanie M, Bart Massi, Samuel R Mathias, Daeyeol Lee, and Jane R Taylor. 2019. "Model-free and model-based influences in addiction-related behaviors." *Biological psychiatry*, 85(11): 936–945.

Grosskurth, Elmar D, Dominik R Bach, Marcos Economides, Quentin JM Huys, and Lisa Holper. 2019. "No substantial change in the balance between model-free and model-based control via training on the two-step task." *PLoS computational biology*, 15(11): e1007443.

Hackel, Leor M, Jeffrey J Berg, Björn R Lindström, and David M Amodio. 2019. "Model-based and model-free social cognition: investigating the role of habit in social attitude formation and choice." *Frontiers in psychology*, 10: 2592.

Haghani, Milad, Michiel CJ Bliemer, John M Rose, Harmen Oppewal, and Emily Lancsar. 2021. "Hypothetical bias in stated choice experiments: Part I. Macro-scale analysis of literature and integrative synthesis of empirical evidence from applied economics, experimental psychology and neuroimaging." *Journal of choice modelling*, 41: 100309.

Hallegatte, Stéphane, Adrien Vogt-Schilb, Julie Rozenberg, Mook Bangalore, and Chloé Beaudet. 2020. "From poverty to disaster and back: A review of the literature." *Economics of Disasters and Climate Change*, 4: 223–247.

Hartston, Heidi. 2012. "The case for compulsive shopping as an addiction." *Journal of psychoactive drugs*, 44(1): 64–67.

Haushofer, Johannes, and Ernst Fehr. 2014. "On the psychology of poverty." *science*, 344(6186): 862–867.

Heo, Suyeon, Yoondo Sung, and Sang Wan Lee. 2021. "Effects of subclinical depression on prefrontal–striatal model-based and model-free learning." *PLOS Computational Biology*, 17(5): e1009003.

Huys, Quentin JM, and Daniel Renz. 2017. "A formal valuation framework for emotions and their control." *Biological Psychiatry*, 82(6): 413–420.

Imataka, George, Ryoichi Sakuta, Akira Maehashi, and Shigemi Yoshihara. 2022. "Current Status of Internet Gaming Disorder (IGD) in Japan: New Lifestyle-Related Disease in Children and Adolescents." *Journal of Clinical Medicine*, 11(15): 4566.

Jolliffe, Darrick, and David P Farrington. 2006. "Development and validation of the Basic Empathy Scale." *Journal of adolescence*, 29(4): 589–611.

Karter, Andrew J, Melissa M Parker, Howard H Moffet, Ameena T Ahmed, Assiamira Ferrara, Jennifer Y Liu, and Joe V Selby. 2004. "Missed appointments and poor glycemic control: an opportunity to identify high-risk diabetic patients." *Medical care*, 110–115.

Katz, Steven J, and Timothy P Hofer. 1994. "Socioeconomic disparities in preventive care persist despite universal coverage: breast and cervical cancer screening in Ontario and the United States." *Jama*, 272(7): 530–534.

Kelley, Ann E, and Kent C Berridge. 2002. "The neuroscience of natural rewards: relevance to addictive drugs." *Journal of neuroscience*, 22(9): 3306–3311.

Keramati, Mehdi, Amir Dezfouli, and Payam Piray. 2011. "Speed/accuracy trade-off between the habitual and the goal-directed processes." *PLoS computational biology*, 7(5): e1002055.

Kim, Dongjae, Geon Yeong Park, John P O' Doherty, and Sang Wan Lee. 2019. "Task complexity interacts with state-space uncertainty in the arbitration between model-based and model-free learning." *Nature communications*, 10(1): 5738.

Kim, Jinhee, Benoit Sorhaindo, and E Thomas Garman. 2006. "Relationship between financial stress and workplace absenteeism of credit counseling clients." *Journal of Family and Economic Issues*, 27: 458–478.

Knutson, Brian, Charles M Adams, Grace W Fong, and Daniel Hommer. 2001. "Anticipation of increasing monetary reward selectively recruits nucleus accumbens." *The Journal of neuroscience*, 21(16): RC159.

Ko, Daijin, and Matthew J Wanat. 2016. "Phasic dopamine transmission reflects initiation vigor and exerted effort in an action-and region-specific manner." *Journal of Neuroscience*, 36(7): 2202–2211.

Kromer, J, T Hummel, D Pietrowski, AS Giani, J Sauter, G Ehninger, AH Schmidt, and I Croy. 2016. "Influence of HLA on human partnership and sexual satisfaction." *Scientific reports*, 6(1): 1–6.

Kurdi, Benedek, Samuel J Gershman, and Mahzarin R Banaji. 2019. "Model-free and model-based learning processes in the updating of explicit and implicit evaluations." *Proceedings of the National Academy of Sciences*, 116(13): 6035–6044.

Labrecque, Lauren I, and George R Milne. 2012. "Exciting red and competent blue: the importance of color in marketing." *Journal of the Academy of Marketing Science*, 40: 711–727.

Landry, Peter, and Ryan Webb. 2021. "Pairwise normalization: A neuroeconomic theory of multi-attribute choice." *Journal of Economic Theory*, 193: 105221.

Lee, Sang Wan, Shinsuke Shimojo, and John P O'Doherty. 2014. "Neural computations underlying arbitration between model-based and model-free learning." *Neuron*, 81(3): 687–699.

Lewis, Robert G, Ermanno Florio, Daniela Punzo, and Emiliana Borrelli. 2021. "The Brain's reward system in health and disease." *Circadian Clock in Brain Health and Disease*, 57–69.

Lieder, Falk, Owen X Chen, Paul M Krueger, and Thomas L Griffiths. 2019. "Cognitive prostheses for goal achievement." *Nature human behaviour*, 3(10): 1096–1106.

Lind, Martina, Mimì Visentini, Timo Mäntylä, and Fabio Del Missier. 2017. "Choice-supportive misremembering: A new taxonomy and review." *Frontiers in psychology*, 8: 2062.

Linnet, Jakob, Kim Mouridsen, Ericka Peterson, Arne Møller, Doris Jeanne Doudet, and Albert Gjedde. 2012. "Striatal dopamine release codes uncertainty in pathological gambling." *Psychiatry Research: Neuroimaging*, 204(1): 55–60.

Li, Norman P, Mark van Vugt, and Stephen M Colarelli. 2018. "The evolutionary mismatch hypothesis: Implications for psychological science." *Current Directions in Psychological Science*, 27(1): 38–44.

Loewenstein, George, Ted O'Donoghue, and Sudeep Bhatia. 2015. "Modeling the interplay between affect and deliberation." *Decision*, 2(2): 55.

Löfgren, Åsa, and Katarina Nordblom. 2020. "A theoretical framework of decision making explaining the mechanisms of nudging." *Journal of Economic Behavior & Organization*, 174: 1–12.

Manus, Melissa B. 2018. "Evolutionary mismatch." *Evolution, medicine, and public health*, 2018(1): 190–191.

Marinelli, Michela, and James E McCutcheon. 2014. "Heterogeneity of dopamine neuron activity across traits and states." *Neuroscience*, 282: 176–197.

Marsh, Abigail A. 2018. "The neuroscience of empathy." *Current opinion in behavioral sciences*, 19: 110–115.

Mathar, David, Annika Wiebe, Deniz Tuzsus, and Jan Peters. 2022. "Erotic cue exposure increases physiological arousal, biases choices towards immediate rewards and attenuates model-based reinforcement learning." *bioRxiv*, 2022–09.

McLoyd, Vonnie C. 1998. "Socioeconomic disadvantage and child development." *American psychologist*, 53(2): 185.

Méjean, Aurélie, Antonin Pottier, Marc Fleurbaey, and Stéphane Zuber. 2020. "Catastrophic climate change, population ethics and intergenerational equity." *Climatic Change*, 163(2): 873–890.

Mikus, Nace, Sebastian Korb, Claudia Massaccesi, Christian Gausterer, Irene Graf, Matthäus Willeit, Christoph Eisenegger, Claus Lamm, Giorgia Silani, and Christoph Mathys. 2022. "Effects of dopamine D2/3 and opioid receptor antagonism on the trade-off between model-based and model-free behaviour in healthy volunteers." *Elife*, 11: e79661.

Moran, Rani, Mehdi Keramati, and Raymond J Dolan. 2021. "Model based planners reflect on their model-free propensities." *PLOS Computational Biology*, 17(1): e1008552.

Moran, Rani, Mehdi Keramati, Peter Dayan, and Raymond J Dolan. 2019. "Retrospective model-based inference guides model-free credit assignment." *Nature communications*, 10(1): 750.

Murtaugh, Paul A, and Michael G Schlax. 2009. "Reproduction and the carbon legacies of individuals." *Global Environmental Change*, 19(1): 14–20.

Neal, Richard D, Debbie A Lawlor, Victoria Allgar, Malcolm Colledge, Shahid Ali, Alan Hassey, Christine Portz, and Andrew Wilson. 2001. "Missed appointments in general practice: retrospective data analysis from four practices." *British journal of general practice*, 51(471): 830–832.

Nederkoorn, C, R Guerrieri, RC Havermans, A Roefs, and A Jansen. 2009. "The interactive effect of hunger and impulsivity on food intake and purchase in a virtual supermarket." *International journal of obesity*, 33(8): 905–912.

Nilashi, Mehrbakhsh, Sarminah Samad, Neda Ahmadi, Ali Ahani, Rabab Ali Abumalloh, Shahla Asadi, Rusli Abdullah, Othman Ibrahim, and Elaheh Yadegaridehkordi. 2020. "Neuromarketing: a review of research and implications for marketing." *Journal of Soft Computing and Decision Support Systems*, 7(2): 23–31.

Nordhaus, William D. 1993. "Optimal greenhouse-gas reductions and tax policy in the" DICE" model." *The American Economic Review*, 83(2): 313–317.

Nutt, David J, Anne Lingford-Hughes, David Erritzoe, and Paul RA Stokes. 2015. "The dopamine theory of addiction: 40 years of highs and lows." *Nature Reviews Neuroscience*, 16(5): 305–312.

Oguchi, Mineki, Yang Li, Yoshie Matsumoto, Toko Kiyonari, Kazuhiko Yamamoto, Shigeki Sugiura, and Masamichi Sakagami. 2023. "Proselfs depend more on model-based than model-free learning in a non-social probabilistic state-transition task." *Scientific Reports*, 13(1): 1419.

O'Sullivan, Sean S, Kit Wu, Marios Politis, Andrew D Lawrence, Andrew H Evans, Subrata K Bose, Atbin Djamshidian, Andrew J Lees, and Paola Piccini. 2011. "Cue-induced striatal dopamine release in Parkinson's disease-associated impulsive-compulsive behaviours." *Brain*, 134(4): 969–978.

Otto, A Ross, Candace M Raio, Alice Chiang, Elizabeth A Phelps, and Nathaniel D Daw. 2013. "Working-memory capacity protects model-based learning from stress." *Proceedings of the National Academy of Sciences*, 110(52): 20941–20946.

Pampel, Fred C, Patrick M Krueger, and Justin T Denney. 2010. "Socioeconomic disparities in health behaviors." *Annual review of sociology*, 36: 349–370.

Panos, John J, and Lisa E Baker. 2010. "An in vivo microdialysis assessment of concurrent MDMA and cocaine administration in Sprague–Dawley rats." *Psychopharmacology*, 209: 95–102.

Pasquereau, Benjamin, and Robert S Turner. 2013. "Limited encoding of effort by dopamine neurons in a cost–benefit trade-off task." *Journal of Neuroscience*, 33(19): 8288–8300.

Pörtner, Hans-O, Debra C Roberts, Helen Adams, Carolina Adler, Paulina Aldunce, Elham Ali, Rawshan Ara Begum, Richard Betts, Rachel Bezner Kerr, Robbert Biesbroek, et al. 2022. *Climate change 2022: Impacts, adaptation and vulnerability.* IPCC Geneva, Switzerland:.

Rauh, Virginia, Srikesh Arunajadai, Megan Horton, Frederica Perera, Lori Hoepner, Dana B Barr, and Robin Whyatt. 2011. "Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide." *Environmental health perspectives*, 119(8): 1196–1201.

Rawls, John. 2004. "A theory of justice." In *Ethics*. 229–234. Routledge.

Reed, Danielle Renee, and Antti Knaapila. 2010. "Genetics of taste and smell: poisons and pleasures." *Progress in molecular biology and translational science*, 94: 213–240.

Sahni, Navdeep S, S Christian Wheeler, and Pradeep Chintagunta. 2018. "Personalization in email marketing: The role of noninformative advertising content." *Marketing Science*, 37(2): 236–258.

Schad, Daniel J, Elisabeth Jünger, Miriam Sebold, Maria Garbusow, Nadine Bernhardt, Amir-Homayoun Javadi, Ulrich S Zimmermann, Michael N Smolka, Andreas Heinz, Michael A Rapp, et al. 2014. "Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning." *Frontiers in Psychology*, 1450.

Schultz, Wolfram. 2015. "Neuronal reward and decision signals: from theories to data." *Physiological reviews*, 95(3): 853–951.

Schwartz, Shalom. 2006. "A theory of cultural value orientations: Explication and applications." *Comparative sociology*, 5(2-3): 137–182.

Shah, Anuj K, Sendhil Mullainathan, and Eldar Shafir. 2012. "Some consequences of having too little." *Science*, 338(6107): 682–685.

Shaw, Steven D, and Richard P Bagozzi. 2018. "The neuropsychology of consumer behavior and marketing." *Consumer Psychology Review*, 1(1): 22–40.

Shefrin, Hersh M, and Richard H Thaler. 1988. "The behavioral life-cycle hypothesis." *Economic inquiry*, 26(4): 609–643.

Solinas, Marcello, Sergi Ferre, Zhi-Bing You, Marzena Karcz-Kubicha, Patrizia Popoli, and Steven R Goldberg. 2002. "Caffeine induces dopamine and glutamate release in the shell of the nucleus accumbens." *Journal of Neuroscience*, 22(15): 6321–6324.

Standage, Dominic, Da-Hui Wang, and Gunnar Blohm. 2014. "Neural dynamics implement a flexible decision bound with a fixed firing rate for choice: a model-based hypothesis." *Frontiers in neuroscience*, 8: 318.

Stanovich, Keith E. 1999. *Who is rational?: Studies of individual differences in reasoning.* Psychology Press.

Stern, Nicholas. 2006. *The economics of climate change: the Stern Review.* Cambridge University Press.

Sternthal, Michelle J, Natalie Slopen, and David R Williams. 2011. "Racial disparities in health: how much does stress really matter? 1." *Du Bois review: social science research on race*, 8(1): 95–113.

Teegarden, Sarah L, Eric J Nestler, and Tracy L Bale. 2008. "ΔFosB-mediated alterations in dopamine signaling are normalized by a palatable high-fat diet." *Biological psychiatry*, 64(11): 941–950.

Thaler, Richard H, and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin.

Thaler, Richard H, and Hersh M Shefrin. 1981. "An economic theory of self-control." *Journal of political Economy*, 89(2): 392–406.

Tol, Richard SJ. 2002. "Welfare specifications and optimal control of climate change: an application of fund." *Energy Economics*, 24(4): 367–376.

Tsur, Yacov, and Amos Zemel. 2009. "Endogenous discounting and climate policy." *Environmental and Resource Economics*, 44: 507–520.

Tversky, Amos, and Daniel Kahneman. 1986. "The framing of decisions and the evaluation of prospects." In *Studies in Logic and the Foundations of Mathematics*. Vol. 114, 503–520. Elsevier.

van Swieten, Maaike MH, Rafal Bogacz, and Sanjay G Manohar. 2021. "Hunger improves reinforcement-driven but not planned action." *Cognitive, Affective, & Behavioral Neuroscience*, 21(6): 1196–1206.

Varazzani, Chiara, Aurore San-Galli, Sophie Gilardeau, and Sebastien Bouret. 2015. "Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys." *Journal of Neuroscience*, 35(20): 7866–7877.

Volkow, Nora D, Gene-Jack Wang, Joanna S Fowler, Dardo Tomasi, Frank Telang, and Ruben Baler. 2010. "Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit." *Bioessays*, 32(9): 748–755.

Voon, Valerie, Thomas B Mole, Paula Banca, Laura Porter, Laurel Morris, Simon Mitchell, Tatyana R Lapa, Judy Karr, Neil A Harrison, Marc N Potenza, et al. 2014. "Neural correlates of sexual cue reactivity in individuals with and without compulsive sexual behaviours." *PloS one*, 9(7): e102419.

Watts, Adreanna TM, and Edward M Bernat. 2018. "Effects of reward context on feedback processing as indexed by time-frequency analysis." *Psychophysiology*, 55(9): e13195.

Wunderlich, Klaus, Peter Smittenaar, and Raymond J Dolan. 2012. "Dopamine enhances model-based over model-free choice behavior." *Neuron*, 75(3): 418–424.