Naiveté and sophistication in dynamic inconsistency

Zsombor Z. Méder, János Flesch, Ronald Peeters

Naiveté and sophistication in dynamic inconsistency

GSBE

Naiveté and sophistication in dynamic inconsistency[*]

Zsombor Z. Méder[†]   János Flesch[‡]   Ronald Peeters[§]

February 2014

Abstract

This paper introduces a general framework for dealing with dynamic inconsistency in the context of Markov decision problems. It carefully decouples and examines concepts that are often entwined in the literature: it distinguishes between the decision maker and its various temporal agents, and between the beliefs and intentions of the agents. Classical examples of naiveté and sophistication are modeled and contrasted based on this new language. We show that naive and sophisticated decision makers can form optimal strategies at each possible history, and provide welfare comparisons for a class of decision problems including procrastination, impulsiveness, underinvestment, binges and indulgence. The creation of a unified formalism to deal with dynamic inconsistency allows for the introduction of a hybrid decision maker, who is naive sometimes, sophisticated at others. Such a hybrid decision maker can be used to model situations where type determination is endogenous. Interestingly, the analysis of hybrid types indicates that self-deception can be optimal.

Keywords: dynamic inconsistency; naiveté; sophistication; Markov decision problem; quasi-hyperbolic discounting

JEL classification: C70, D11, D91

[†]Department of Economics, Maastricht University, The Netherlands. E: z.meder@maastrichtuniversity.nl. Corresponding author

[‡]Department of Quantitative Economics, Maastricht University, The Netherlands. E: j.flesch@maastrichtuniversity.nl

[§]Department of Economics, Maastricht University, The Netherlands. E: r.peeters@maastrichtuniversity.nl

1 Introduction

Imagine that you are sitting with a few friends, drinking beer. You have just finished your second glass, and your friends want to order a new round. You think to yourself: "well, I could deal with one or two more, but then I really should go home." However, from previous experience, you are also acutely aware that after your third beer your mindset is likely to change: you will start fooling yourself, repeating over and over in the course of the evening: "just one more beer, and then I am *really* going home". This would lead to an undesirable outcome, getting drunk and having a hangover the next morning. So you wisely leave your friends after just your second beer. What is happening here? The framework that we propose in this paper makes it possible to model this and similar scenarios.[1]

Traditionally, there are three main ways to portray decision makers under time inconsistency. The first one regards decision makers as naifs (Akerlof, 1991; O'Donoghue and Rabin, 1999b), the second attributes sophistication to them (Laibson, 1997; Fischer, 1999; Harris and Laibson, 2001), while the third argues that even resolute behavior is possible (McClennen, 1990). A common assumption of these models in the classic papers on dynamic inconsistency is regarding the decision maker as falling entirely into one of the above three categories, treating his type as exogeneously given. More recently, mainly building on the work of O'Donoghue and Rabin (2001), hybrid decision makers have been considered, in models of so-called "partial naiveté". However, these models still treat this type as exogeneous to the decision problem.

One way to interpret our example above is that you are sophisticated after finishing the first two beers, but as you drink more, you expect to become naive later on. The story indicates that in certain situations, a decision maker could cause his type to change; moreover, he might even be able to reason about such changes. Any perspective that assumes a fixed type is unable to capture such situations. In order to fix this shortcoming, we attempt a general interpretation of naiveté and sophistication for dynamic inconsistency. The language we develop allows the introduction of hybrid-type decision makers such as the one in our example.

After reviewing the relevant literature, we start building a formalism that allows for precise definitions of the two most commonly discussed types of decision makers, naifs and sophisticates. We work in discrete time, assuming that the situation of the decision maker can be captured as a Markov decision problem. Adopting the terminology of multiple self models, we distinguish between the agent level and the level of the decision maker. Starting on the agent level, we introduce the notion of a strategy, which contains information about the

---

[1]The model of this situation is presented in Section 7.

intentions and beliefs of the individual agent at one particular history. Next, the properties of strategies (coherence, stationarity, consistency) are discussed. Moving to the level of the decision maker (i.e., the collection of all agents), we define the concept of a *frame*, and consider its properties. In two theorems, we link the properties of strategies to the properties of frames. By the end of this foundational section, we possess a useful tool for modelling dynamic inconsistency in discrete time.

After clarifying our assumptions on utility functions, we define and introduce the two pure types of decision making, naiveté and sophistication. We provide existence results for optimal frames of both types, and discuss various properties of such frames. In the main text body, we then expand the framework to include decision makers with a hybrid type, also providing an existence theorem, and discuss two examples of hybrid decision making in detail. In the concluding section, we point towards further extensions of the model. The appendix gives a minimalist summary of classical decision problems of the literature on dynamic inconsistency, and compares naively and sophisticatedly optimal frames for these problems. It also serves as an illustration for the applicability of our approach.

The contributions of this paper are thus threefold. First, it provides an effective language and toolbox for problems of dynamic inconsistency. In particular, the representation of the decision maker through a frame and the theorems relating consistency and stationarity should prove useful (Section 3.6). Second, it provides definitions for naiveté and sophistication, and shows the existence of naively and sophisticatedly optimal frames. Third, it introduces hybrid naive-sophisticated types, expanding the scope of dynamic inconsistency models. As a corollary, the analysis of hybrid types shows that self-deception can be optimal.

2 Related literature

The standard overview on time discounting, various discount functions, and the empirical evidence on time preference is provided by Frederick et al. (2002); while Soman et al. (2005) summarize the various avenues of research in this area. The most widely studied model of non-exponential time preferences is hyperbolic discounting (Phelps and Pollak, 1968; Loewenstein and Thaler, 1989; Loewenstein and Prelec, 1992; Laibson, 1994, 1997). Consumption derived from hyperbolic discount functions can exhibit many anomalies (Harris and Laibson, 2001). One strand of the literature attempts to explain or rationalize hyperbolic models with reference through uncertain lifetime length (Bommier, 2006; Halevy, 2008) or uncertainty about future discount rates (Azfar, 1999). On the other hand, the validity of hyperbolic (or "beta-delta") discounting in explaining intertemporal choice has been criticized by some (Rubinstein, 2003; Benhabib et al., 2010; Andersen et al., 2011). For decision problems in

discrete time, quasi-hyperbolic discounting became the most prominent model following the work of Laibson (1997). However, experimental evidence with regards to the prevalence of this particular form of discount function is not conclusive (Benhabib et al., 2010).

Our paper does not assume any particular form of discounting when considering deviations from the classic, exponential discounted utility model of Samuelson (1937) (see Section 4). However, for ease of presentation, all of our examples will make use of quasi-hyperbolic discounting. In fact, it turns out that most phenomena classically studied in the area of dynamic inconsistency can be modelled as rather simple Markov decision problems. In particular, when analyzing and comparing decision makers that are purely naive or sophisticated, we will deal with the problems of impulsiveness, procrastination, addiction, underinvestment and indulgence (see the appendix).

We should note that in our paper, these concepts have a more narrow meaning than in the literature. For example, Akerlof (1991) talks about substance abuse and savings under the heading of "procrastination", while Ainslie (1975) gives the very same phenomena the label "impulsiveness". To contain these fluctuations of terminology, we try to attach these concepts to a particular set of Markov decision problems.

There is ample evidence that many individuals have issues with procrastination. Fischer (1999) compares a variety of approaches to model this behavior, including the possibility to derive it from time-consistent preferences. In general, individuals use various internal and external commitment devices to tackle the problem of procrastination; however, when these commitment devices are set voluntarily, they are not necessarily optimal (Ariely and Wertenbroch, 2002). For an overview of the connections between procrastination and (social) psychology on procrastination, see Akerlof (1991) and Steel (2007). We interpret procrastination problems as involving a single task that would induce benefits in the longer, but not in the shorter term.

Impulsiveness has already been studied by behavioral psychologists several decades ago (Ainslie, 1975). Impulse control is an essential skill for adaptation: Ainslie (1974) reports that even pigeons are able to exercise impulse control in order to receive larger rewards later. In our framework, impulsive actions can only be executed once, and have positive short-term, but large negative long-term consequences.

Similar to procrastination, it is common for people to repeatedly fail to choose an action that would bring long-term benefits. DellaVigna and Malmendier (2004) report that most people buy costly gym passes that they subsequently underuse. Laibson et al. (1998) fit a model of (naive) households with quasi-hyperbolic discounting to life-cycle consumption and saving patterns, providing ample evidence for wide undersaving for retirement; Angeletos et al. (2001) find similar results. Our notion of underinvestment involves a *repeated* decision

between a large long- and a small short-term benefit.

There is a large number of studies analyzing addiction from an economic perspective, most notably relating to the rational addiction hypothesis (Becker and Murphy, 1988). More relevant to our paper is the work of Ainslie and Monterosso (2003), who overview the links between addiction and hyperbolic discounting. Because the notion of addiction is quite laden with physiological connotations, we will instead talk about "binges", i.e., actions that are repeatedly available to the decision maker that bring him harm in the longer term. Gruber and Kőszegi (2001) consider the link between self-control and addiction.

Indulgence is a particular type of timing problem, where the issue is when to consume a certain good. We adopt this model directly from O'Donoghue and Rabin (1999a), as the simplest decision problem we are aware of where the welfare ranking between a naive and a sophisticated decision maker favors naiveté.

Modelling approaches to dynamic inconsistency come in two varieties (Asheim, 2007). Dual-self planner-doer models bear a close analogy to principal-agent models (Thaler and Shefrin, 1981). A (single) *planner*, endowed with dynamically consistent preferences formulates plans; a present-biased *doer* can execute them or deviate from them. The conceptual background of planner-doer models is multifold: they sometimes rely on the hot-cold empathy gap (Loewenstein, 2005), on recent findings of neuroscience, or even the Freudian distinction between the id and the ego. Fudenberg and Levine (2006) argues in favor of dual-self model as being more simple analytically, more in line with findings in neuroscience, and nevertheless being able to explain a large number of empirical phenomena. An important advantage of this approach is that welfare comparisons are more straightforward: the preferences of the planner are generally adopted to be normatively relevant. Recent models allow for the planner to learn about the doer's type through costly experimentation (Ali, 2011), or can include self-control (Fudenberg and Levine, 2012; Benabou and Pycia, 2002).

Despite their advantages, dual-self models make a substantive normative assumption: the long-run preferences of an individual can be identified. There is no clear a priori reason why this should be the case. In this paper, we will avoid this assumption, and instead adopt the multiple-self approach which maintains the ontological equivalence between the various conditions in which an individual can find himself.[2]

Our primary focus is on naive and sophisticated decision makers, and hybrid types. Multiple-self models have used pure types of decision makers in a variety of settings. For instance, Fischer (1999); Laibson (1994, 1997); Angeletos et al. (2001) work with sophisticates, while Akerlof (1991) and O'Donoghue and Rabin (1999b) assume naiveté. In a generial equilibrium setting, Herings and Rohde (2006, 2008) deal with both naifs and sophisticates.

---

[2]Bach and Heilmann (2011) link multiple-self models to the philosophical literature on personal identity.

For us, a crucial precursor paper is O'Donoghue and Rabin (1999a), who systematically compare naiveté and sophistication for the case of quasi-hyperbolic discounting, with so-called immediate costs and rewards.

The first steps towards hybrid naivé-sophisticated types have been taken by O'Donoghue and Rabin (2001). They define partial naiveté in the context of quasi-hyperbolic ($\beta - \delta$-)discounting. While naifs think their β is 1, the actual β is fully known to sophisticates, while partially naive agents think they have a β that is larger than their actual one. Partially naive agents thus entertain false beliefs about the future (just like naifs). This approach has proven its fruitfulness especially in the contract design literature (Gilpatric, 2003; DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006); while DellaVigna and Malmendier (2006) focus on a monopolistic firm facing a mixed population of consumers. Heidhues and Kőszegi (2009) generalize the distributional assumptions on beliefs, but still in the context of quasi-hyperbolic discounting.

An important aspect of the O'Donoghue and Rabin (2001) approach is that – compared to sophistication – *any* degree of naiveté can generate arbitrarily large losses in efficiency for a decision maker. Thus the limit of partial naiveté (as perceived present-biasedness approaches the real parameter) is not sophistication.

There were a few other attempts to treat hybrid decision makers. In Asheim (2007), agents have a perceived preference persistence, i.e., a probability (between 0 and 1) with which they think their preferences will be identical in the next period. However, this belief is always incorrect, as their preferences will change with probability 1. In Jehiel and Lilico (2010) agents endowed with exponential discounting have access to information about a number of future periods ("foresight"). They find that improving the length of foresight always improves welfare.

This paper expands on the existing literature by providing the foundations of a hybrid model that is independent of the quasi-hyperbolic assumption of O'Donoghue and Rabin (1999a), and where type determination is an endogenous part of the decision problem, as in our motivating example.

3   Basic concepts

In this section, we introduce our framework and notations. Standard definitions are provided for the notions of "decision problem" and "history". However, before introducing strategies, we argue in Section 3.3 for a new way of defining strategies that includes both beliefs and intentions. We then proceed by defining strategies and frames, which correspond to two different levels of analysis. Towards the end of the section, we present some results on the

relationship between consistency and stationarity.

3.1 Markov decision problem

We start with a decision maker facing a finite Markov decision problem on an infinite horizon.[3]

Definition 1. A *finite Markov decision problem* is given by:

- the set of time periods $T = \{0, 1, 2, \dots\}$;

- a finite set of states Ω, with $\bar{\omega} \in \Omega$ as the initial state;

- a finite and non-empty set of pure actions A_ω that the decision maker can choose from in state ω;

- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action in state ω;

- transition probabilities $m_\omega : A_\omega \to \Delta(\Omega)$, with $m_\omega(\omega'|a_\omega)$ denoting the probability to transit from state ω to state ω' when action a_ω is chosen.

This definition excludes the possibility of randomization over actions, although in some cases such an extension is needed to obtain existence of a stationary optimal strategy (see Section A.2). Nevertheless, in order to keep the presentation simpler, for the moment we only consider pure actions.

3.2 History

To capture all the informational aspects on which an action choice can be conditioned, we introduce the notion of a history:

Definition 2. A *history* h has the form $h = \left(\omega_0, a_{\omega_0}, \dots, \omega_{t-1}, a_{\omega_{t-1}}, \omega_t\right)$, with:

- $\omega_i \in \Omega$, for $i \in \{0, 1, \dots, t\}$, and $\omega_0 = \bar{\omega}$;

- $a_{\omega_i} \in A_{\omega_i}$, for $i \in \{0, 1, \dots, t-1\}$;

- $m_{\omega_i}(\omega_{i+1}|a_{\omega_i}) > 0$, for $i \in \{0, 1, \dots, t-1\}$.

The length of h or current time at h is denoted by $t = t(h)$, and the function $\omega(h) = \omega_{t(h)}$ indicates the current or end state at history h. We use H to refer to the set of all histories.

---

[3]The latter is not a restrictive requirement, since it is easy to rewrite a decision problem on a finite horizon to one on an infinite horizon.

If history h' begins with h, we say that h' *succeeds* h, or equivalently, that h *precedes* h'; and denote this with $h \preccurlyeq h'$, or equivalently, with $h' \succcurlyeq h$.[4] The subset of H that consists of all histories that succeed h is denoted by $H^{\succcurlyeq h}$:

$$H^{\succcurlyeq h} = \{h' \in H \mid h' \succcurlyeq h\}.$$

We refer to $h_0 = (\omega_0)$ as the "root history". To shorten notation, when specifying a history, we sometimes omit the commas separating states and actions, and also the enclosing parentheses. Thus, history $h = (\omega_0, a_{\omega_0}, \omega_1)$ will be occasionally written as $h = \omega_0 a_{\omega_0} \omega_1$.

3.3 Conceptual foundations

We have to make a fundamental distinction between these multiple selves, which we dub "agents", and the notion of a "decision maker". The fundamental entities in our model are *agents*: they are the ones with the ultimate power of choosing an action *and* executing it in a certain state. We assume a one-to-one correspondence between histories and agents, since the information available at each of these histories is different, and the action choice might be contingent on such information. We refer to agents as being "at" a history; they have preferences, and can relate to the future. Our agents form beliefs and intentions about the future, and have the ability to choose an action in the present.[5] Thus, our notion of a "strategy", as defined in Section 3.4, slightly departs from the standard definition, as it keeps this separation of beliefs and intentions.

To refer to the unity and shared aspects of all temporal selves, we use the notion of *decision maker*. These shared aspects are threefold: first, they refer to the fact that the underlying decision problem is fundamentally the same. Second, they suggest that the preferences of various temporal selves are similar, in the sense that a common functional form can be used to represent them (see Section 4). Third, it alludes to the elusive issue of *personal identity*, the fact that the connections between various temporal selves is substantively more intimate than the similarities between distinct individuals. We use the notion of a decision maker as comprising these three aspects, but carrying no normative weight: the decision maker is not in conflict with any of its particular manifestations.

Returning our focus to agents, we note that although they are temporally distinct, they are related strategically, since the well-being[6] of each agent usually depends on the actions taken by other agents. In the most general case, the well-being of each agent can be decomposed into three components: well-being generated by past actions, current actions and future actions.

---

[4]Obviously, $h' \preccurlyeq h$ & $h \preccurlyeq h' \Leftrightarrow h = h'$.

[5]Conceptually, we follow Cowen (1991), who identifies a self ''with a set of preferences linked to certain cognitive and volitional capacities."

[6]The notion of well-being that we use in this section is close to the meaning of the term "utility", but stripped of its technical connotations.

Past actions are encoded in the history. We keep the general assumption that past actions have no effect on the well-being generated by current and future actions of the agents. In other words, we stick to the notion of "bygones are bygones". This might seem an obvious remark, but one could imagine otherwise. For example, if one is thinking about hiring a private investigator to find out whether his wife has been cheating on him, it is the utility generated by her and his own past actions (e.g. not paying sufficient attention to her) that is potentially being re-evaluated.

As for the present, it is ordinarily assumed that the agent has full control over at least his current action. We stick to this assumption in the current paper.[7]

The third determinant of well-being can be the future actions of the decision maker. Using the distinction between experienced and decision utility (Kahneman et al., 1997), we can delimit two senses of "expected utility from taking action a". Let us disregard the immediate payoff for taking the action, and consider only future payoff. In one sense, the phrase could mean "experienced utility from expectation", i.e., utility that *is actually experienced* by the agent due to expecting a certain stream of future payoffs. Think of a student that decides to study for an upcoming exam instead of watching his favorite TV show. He might, in fact, *already* enjoy the benefits of the decision to study (he is already less anxious for the exam, maybe he relishes the idea that he is doing "the right thing" etc.). The other sense of "expected utility" could be rendered as "the expected present value of various streams of payoffs" that the agent's current decision can lead to. In this sense, the agent is not experiencing any actual change in utility by choosing one or other course of action; he is merely able to calculate with these future payoffs. This distinction between the two senses of expected utility will be used in Section 4, and in our conceptual interpretations of naiveté and sophistication in subsequent sections.

3.4 Intentions, beliefs, strategy

We attempt to give a full description for the two most prevalent decision maker types (naifs and sophisticates) and the hybrid types that we introduce later in the paper. For this purpose, we define a strategy as having three components: the current action, the intended future actions, and the belief on what future agents will in fact do. There is no special reason for assuming that the latter two coincide for future actions, although with our definitions, they coincide for pure, but not for hybrid decision makers. To simplify notation, we reduce this triadic framework to just intentions and beliefs, and assume that for the current action, these

---

[7]To see that this choice is not so obvious, see Jehiel and Lilico (2010). Similarly, Elster's interpretation of the Ulysses story is an example of a model where control over current action is essentially eliminated (Elster, 1979).

two have to coincide: no agent can be wrong about which action he takes, and each agent takes the action that he intends to.

The basic building blocks of our model are all functions from the set of histories that succeed the agent to the set of available actions at those histories.

Definition 3. The *intentions* of an agent at history $\bar{h}$ assign an intended action to each history that succeeds the present:

$$i^{\bar{h}} : h \in H^{\succeq \bar{h}} \mapsto A_{\omega(h)}.$$

Definition 4. *The* beliefs *of an agent at history $\bar{h}$ assign an action to each history that succeeds the present:*

$$b^{\bar{h}} : h \in H^{\succeq \bar{h}} \mapsto A_{\omega(h)}.$$

Note that intentions and beliefs are defined at all succeeding histories, even at those that the agent does not intend to reach or believes will not be reached.

We now proceed to define strategies as the pair of intentions and beliefs for an agent. This practice is not common: beliefs are usually separated from the strategy. However, it turns out to be very convenient to include them as part of the strategy.

Definition 5. A *strategy* of an agent at history $\bar{h}$ is a pair of intentions and beliefs for that agent, with the added property that the belief and intention for the current action coincide:

$$s^{\bar{h}} = \left(i^{\bar{h}}, b^{\bar{h}} \right), \text{with } i^{\bar{h}} \left(\bar{h} \right) = b^{\bar{h}} \left(\bar{h} \right).$$

The set of all strategies for this agent is denoted by $S^{\bar{h}}$.

For an agent at $\bar{h}$, $i^{\bar{h}}(h)$ refers to the intention, while $b^{\bar{h}}(h)$ refers to the belief component of the strategy $s^{\bar{h}}$ about the future agent at h. For example, $b^{\bar{h}}(h) = a$ should be read as such: "The agent at $\bar{h}$ who holds strategy $s^{\bar{h}}$ believes the agent at h will choose action a."

We emphasize that our definition of a strategy in terms of an intention-belief pair is not standard. In fact, the most common practice is to conflate the two concepts. On the other hand, in epistemic game theory, beliefs are strictly seperated from actions. If our usage seems awkward, it might be convenient for the reader to think of the term "strategy" simply as an abbreviation for "intention-belief pair".

The above definitions bring us close to a full epistemic characterization of intra-personal decision making. The full epistemic framework should include not only beliefs about the actions of future agents, but also beliefs about future agent's beliefs about future agent's beliefs etc. Moreover, it should also include beliefs about intentions, beliefs about beliefs about intentions etc. It is more controversial whether it should include intentions about

intentions, intentions about beliefs,[8] intentions about beliefs about intentions, or any sequence of intention- and belief-operators, for that matter. Our current goal is just to provide an adequate characterization of naiveté and sophistication, and we can avoid going into such details.

We can now define two properties of strategies, stationarity and coherence, as well as a relation over the set of strategies, consistency.

Definition 6. The intentions (or beliefs) of an agent at $\bar{h}$ are *stationary* whenever the intended (believed) actions depend only on the end-state. Formally, $i^{\bar{h}}$ or $b^{\bar{h}}$ is called stationary if, for all $h, h' \in H^{\succcurlyeq \bar{h}}$ with $\omega(h) = \omega(h')$, we have $i^{\bar{h}}(h) = i^{\bar{h}}(h')$ or respectively, $b^{\bar{h}}(h) = b^{\bar{h}}(h')$. A strategy $s^{\bar{h}}$ is stationary if both its constituent intentions $i^{\bar{h}}$ and beliefs $b^{\bar{h}}$ are stationary.

For example, if each day of the week can be modeled as a single state, the strategy of an agent who intends and believes eating in a restaurant every second Saturday, but staying home on every other one is not stationary.

Definition 7. An agent at $\bar{h}$ is said to hold a *coherent* strategy, if his intention and beliefs about future actions coincide for all future histories. Formally, a strategy $s^{\bar{h}} = \left(i^{\bar{h}}, b^{\bar{h}} \right)$ of an agent at $\bar{h}$ is coherent if $i^{\bar{h}}(h) = b^{\bar{h}}(h)$ for all $h \in H^{\succcurlyeq \bar{h}}$.

For example, a strategy of an agent who intends to stop drinking, but believes he will be unable to do so is not coherent.

Definition 8. The strategies of two agents at h and h' are said to be *consistent* if they assign the same intentions and beliefs to each history that succeeds both agents, i.e., s^h and $s^{h'}$ are consistent, if $s^h(h'') = s^{h'}(h'')$ for all $h'' \in H^{\succcurlyeq h} \cap H^{\succcurlyeq h'}$.[9]

For example, a strategy formulated yesterday which intended eating apples for today as desert and a strategy formulated today that intends eating cookies instead are not consistent.

Whereas coherence concerns the relationship between the intentions and beliefs of the same strategy, i.e., belonging to one agent, consistency compares strategies of two distinct agents. In other words, coherence is an intrinsic property, whereas consistency is an extrinsic (relational) property of a strategy.

A natural question is whether consistency of strategies is transitive, i.e., whether if s^h and $s^{h'}$ are consistent, and $s^{h'}$ and $s^{h''}$ are also consistent, for some $h \preccurlyeq h' \preccurlyeq h''$, implies that s^h

---

[8]The toxin puzzle seems to indicate that there are some scenarios in which an agent might have an intention to form a future intentions i, but would be unable to ever form i. Also, aware of this impossibility, he would not believe he will form i (Kavka, 1983).

[9]So, if two strategies are defined at histories that neither succeed nor preceed each other, then they are consistent, as there are no histories that succeed both.

and $s^{h''}$ are also consistent. However, without this constraint, consistency is not transitive in general – it is not even transitive within the set of stationary strategies. To see this, take the decision problem in Figure 1.[10] We construct three stationary strategies s^{h_ρ}, s^{h_σ}, and s^{h_τ} such that s^{h_ρ}, s^{h_σ} are consistent, as well as s^{h_σ}, s^{h_τ}, but s^{h_ρ}, s^{h_τ} are not consistent. Fix $h_\rho = (\rho)$, $h_\sigma = (\rho, A, \sigma)$ and $h_\tau = (\rho, B, \tau)$. Also, let $s^{h_\rho}(h_\rho) = (A, A)$, $s^{h_\rho}(h) = (C, C)$ if $\omega(h) = \sigma$, and $s^{h_\rho}(h) = (E, E)$ if $\omega(h) = \tau$. Intuitively, s^{h_ρ} means: "I choose A, believe and intend C in state ρ, and believe and intend E in state τ." Define two other strategies through $s^{h_\sigma}(h) = (C, C)$ for all $h \succcurlyeq h_\sigma$ ("do C in σ"), and $s^{h_\tau}(h) = (F, F)$ for all $h \succcurlyeq h_\tau$ ("do F in τ"). All of these strategies are stationary. Clearly, s^{h_ρ} and s^{h_σ} are consistent, since they both require the decision maker to choose C in state σ, and after history h_σ no state state other than σ is reachable. Next, s^{h_σ} and s^{h_τ} are consistent, since histories h_σ and h_τ neither succeed, nor precede each other. But s^{h_ρ} and s^{h_τ} are not consistent, as they assign different actions to the state τ. This shows that consistency of strategies is not transitive on the set of stationary strategies.
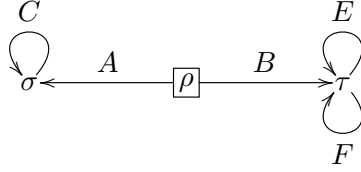
Figure 1: Stationary strategies with intransitive consistency.

3.5   Truncation

The following definitions of "truncation", although they look like technicalities, are necessary for the definition of a stationary frame. Truncation formalizes the idea of "bygones are bygones", and chips away an initial segment of the history.

Definition 9. Take any history $h = (\omega_0, a_{\omega_0}, \dots, \omega_{t(h)})$. The *truncation* operator $\dashv_k$, defined for any $k \leq t(h)$, removes the first k pairs of this sequence, so $\dashv_k h = (\omega_k, a_{\omega_k}, \dots, \omega_{t(h)})$. For a set of histories $H' \subseteq H$, we refer to the set of k-truncated histories by $\dashv_k H'$.

A truncated strategy applies this idea to strategies: it is defined for an agent that "forgot" (or disregards) all of its past; future histories obviously do not include descriptions of forgotten

---

[10]We use figures like this one to represent decision problems. States are denoted by greek characters; in this decision problem, we have states ρ, σ, τ, with $\boxed{\rho}$ indicating the initial state. For actions, we use roman capitals – in this case, we have action A, B (available in state ρ), C (the only action available in σ), and E, F (in τ). Whenever – unlike in this example – payoffs are important, the payoff associated with an action is written in the form $A|3$, meaning that choosing action A generates a payoff of 3. Our examples involve only deterministic transition probabilities, and the transitions associated with each state are represented by arrows. For example, for the decision problem in Figure 1 choosing action E in state τ implies that the next state will be τ with probability 1.

segments of the past anymore. Thus, the truncated strategy of an agent at $\bar{h}$ will be defined on the set $\dashv_{t(\bar{h})} H^{\succcurlyeq \bar{h}}$.

Definition 10. For any strategy $s^{\bar{h}}$, the *truncated* strategy $\dashv s^{\bar{h}} : h \in \dashv_{t(\bar{h})} H^{\succcurlyeq \bar{h}} \mapsto A_{\omega(h)}$ denotes the function for which $\dashv s^{\bar{h}} \left(\dashv_{t(\bar{h})} h \right) = s^{\bar{h}}(h)$, for all $h \in H^{\succcurlyeq \bar{h}}$.

We note that while truncated histories are defined for truncations of arbitrary length $(\dashv_k)$, for strategies we only need truncations of length $t(\bar{h})$ for a strategy $s^{\bar{h}}$ of an agent at history $\bar{h}$.

To see this definition at work, think of the decision to stop smoking on the first day of the next month. Take an agent who resolves on July 24th: "I will stop smoking from August 1st" and then fails. On August 24th, he makes another decision: "I will stop smoking from September 1st". If we interpret each statement as a strategy, it is easy to see that they are not consistent: they prescribe different smoking behavior for instance, for August 28th – the first strategy forbids it, while the second allows it. However, there is an intuitive sense in which they are very similar. Indeed, they map them into the same resolve that uses indexicals instead of precise dates: "I can smoke for one more week, and then I will stop".[11] Truncating the present history highlights this similarity by getting rid of the past. In our example, the original strategies are not identical or consistent; but their truncated versions are identical.

3.6 Frames

We now move from the agent level to the level of the decision maker. Since there is no *a priori* reason for the agents to have consistent strategies, different agents can form different intentions and entertain different beliefs about any certain future agent. To have an "external" overview of all agents, we introduce the concept of a frame. In our terminology, a frame is an auxiliary tool for representing the strategies of all possible agents, and not something that is intentionally put together by the decision maker. Whereas each agent chooses a strategy, the decision maker does not choose a frame. Instead, a frame contains a full description of the intentions and beliefs in all contingencies, i.e., at all histories.

Definition 11. A *frame* is a function $f : h \in H \mapsto S^h$. Intuitively, a frame assigns a strategy to each agent.

Figure 2 shows an extremely simple decision problem, for which an example of a frame is represented in Table 1.[12] Each entry is a pair of A's and B's, an intended action and a belief

---

[11]Actually, the difference between specifying a future consumption period by a *calendar date* or through its *time distance* from the present has already been noticed by Strotz (1956). This difference is experimentally explored by Read et al. (2005), who find that subjects only exhibit hyperbolic discounting when future periods are identified via their temporal distance.

[12]For simplicity, we omit parentheses and commas when representing histories in tables.

about an action. Each row corresponds to a strategy for an agent at $\bar{h}$, defining an intention and a belief for each history that succeeds $\bar{h}$. For example, in our table, the entry AB for row $\bar{h} = \rho A \rho$ and column $h = \rho A \rho A \rho$ should be interpreted as: the agent at $\rho A \rho$ intends to choose action A at history $\rho A \rho A \rho$, while believing the agent at $\rho A \rho A \rho$ will, in fact, choose action B. The whole frame thus specifies the intentions and beliefs of all agents over all other (present and future) agents. Our definition of a strategy ensures that the diagonal of the table contains identical actions.
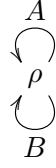
Figure 2: A basic decision problem.

		$h \in H^{\succeq \bar{h}}$					
		ρ	$\rho A \rho$	$\rho B \rho$	$\rho A \rho A \rho$	$\rho A \rho B \rho$	$\ldots$
$\bar{h}$	ρ	BB	AB	AB	AA	BB	$\ldots$
	$\rho A \rho$	$-$	BB	$-$	AB	BA	$\ldots$
	$\rho B \rho$	$-$	$-$	AA	$-$	$-$	$\ldots$
	$\rho A \rho A \rho$	$-$	$-$	$-$	BB	$-$	$\ldots$
	$\rho A \rho B \rho$	$-$	$-$	$-$	$-$	AA	$\ldots$
	$\ldots$	$-$	$-$	$-$	$-$	$-$	$\ldots$

Table 1: An example of a frame for the decision problem in Figure 2.

We now proceed to introduce three properties of frames. Our definition of stationarity makes use of the truncation operator defined above.

Definition 12. A frame f is said to be *stationary*, if only the end-state matters when assigning strategies to histories, i.e., if, for any histories h and h', if $\omega(h) = \omega(h')$, then $\dashv f(h) = \dashv f(h')$.

Stationarity of a frame is different from the stationarity of the strategies involved. For the decision problem in Figure 2, Table 2 offers a non-stationary frame of stationary strategies. To check for this, what needs to be verified is that each row represents a stationary strategy. As there is only one state for this decision problem, this means that in each row, we should see the same intention-belief pair, which is indeed the case. Thus, Table 2 shows a frame of stationary strategies. The frame itself however is not stationary: by truncating the strategies in the first and second row, we get a different strategy.

On the other hand, Table 3 shows a stationary frame of non-stationary strategies. It is easy to see that it is a frame of non-stationary strategies, as each row represents one strategy,

14

which are not stationary – for instance, $s^\rho(\rho) = (B,B) \neq (A,A) = s^\rho(\rho A\rho)$. The frame itself is stationary, which can be checked by comparing the rows. As there is only one state, we need to compare the truncations of all strategies. It can be read from Table 3 that by "forgetting the past", we get the same strategy in each row, namely: the agents picks action B right away, and intends to choose action A, and believes he will do so always in the future.

		$h \in H^{\succeq \bar{h}}$					
		ρ	$\rho A\rho$	$\rho B\rho$	$\rho A\rho A\rho$	$\rho A\rho B\rho$	$\ldots$
	ρ	AA	AA	AA	AA	AA	AA
	$\rho A\rho$	$-$	BB	$-$	BB	BB	BB
$\bar{h}$	$\rho B\rho$	$-$	$-$	AA	$-$	$-$	AA
	$\rho A\rho A\rho$	$-$	$-$	$-$	AA	$-$	AA
	$\rho A\rho B\rho$	$-$	$-$	$-$	$-$	BB	BB
	$\ldots$	$-$	$-$	$-$	$-$	$-$	$\ldots$

Table 2: Non-stationary frame of stationary strategies.

		$h \in H^{\succeq \bar{h}}$					
		ρ	$\rho A\rho$	$\rho B\rho$	$\rho A\rho A\rho$	$\rho A\rho B\rho$	$\ldots$
	ρ	BB	AA	AA	AA	AA	AA
	$\rho A\rho$	$-$	BB	$-$	AA	AA	AA
$\bar{h}$	$\rho B\rho$	$-$	$-$	BB	$-$	$-$	AA
	$\rho A\rho A\rho$	$-$	$-$	$-$	BB	$-$	AA
	$\rho A\rho B\rho$	$-$	$-$	$-$	$-$	BB	AA
	$\ldots$	$-$	$-$	$-$	$-$	$-$	$\ldots$

Table 3: Stationary frame of non-stationary strategies.

For a concrete example of a stationary frame of non-stationary strategies, think of the decision maker who, waking up every day, decides to take *just one more* shot of heroin, and intends (and believes) to quit the next day.

Next, we define a consistent frame. The intuitive idea is that a frame is consistent if no deviation can be expected from previous intentions and beliefs.

Definition 13. A frame f is said to be *consistent* if the strategies $f(h)$ and $f(h')$ assigned to any two histories h and h' are consistent.

Consistency is a very strong notion: a consistent decision maker – at whichever history he is contemplating it – would never change his mind about any intention-belief pair, and thus actual choice. An example would be a heroin user who goes cold turkey immediately and definitely, never ever restarting his substance use. If he relapses, he will not be consistent anymore: his choice to quit for good implies that at the time of quitting his intentions and

15

beliefs for the future history at which he relapses do not match his intentions and beliefs at the point of relapse for the (then-)current history.

According to this definition, if f is a consistent frame, then we get $f(h)\,(h'') = f\,(h')\,(h'')$ whenever $h'' \in H^{\succcurlyeq h} \cap H^{\succcurlyeq h'}$, and either $h' \in H^{\succcurlyeq h}$, or $h \in H^{\succcurlyeq h'}$. Note that, since for a consistent frame, $f(h)(h) = f(h')(h)$ for all h' and $h \in H^{\succcurlyeq h'}$; and our requirement for strategies that $i^h(h) = b^h(h)$, a consistent frame is necessarily made up by coherent strategies. This implies that a choice of an action for all histories uniquely determines a consistent frame. Similarly, a choice of an action for all states uniquely determines a consistent frame of stationary strategies.

Theorem 1. *A consistent, stationary frame consists of stationary strategies.*

Proof. Take any histories h, h' and h'' for which $\omega(h') = \omega(h'')$. We have to show that $f(h)\,(h') = f(h)\,(h'')$. For this, see that:

$$f(h)\,(h') = f\,(h')\,(h') = \dashv f\,(h')\,\big(\dashv_{t(h')} h'\big) = \dashv f\,(h'')\,\big(\dashv_{t(h'')} h''\big) = f\,(h'')\,(h'') = f(h)\,(h'').$$

For the respective equations, we use, in order, consistency, definition of truncation, stationarity of the frame, definition of truncation and consistency again. $\square$

Theorem 2. *A consistent frame of stationary strategies is a stationary frame.*

Proof. Take histories h, h', h'', h''' with $\omega(h) = \omega(h')$. We have to show that

$$(\dashv f(h))\,(\dashv_{t(h)} h'') = (\dashv f(h'))\,(\dashv_{t(h')} h''')$$

if $\dashv_{t(h)} h'' = \dashv_{t(h')} h'''$. Note that such h'', h''' exist, because the end-state in h and h' is identical. Using the fact that $\omega(h'') = \omega(\dashv_{t(h)} h'') = \omega(\dashv_{t(h')} h''') = \omega(h''')$, we get:

$$(\dashv f(h))\,(\dashv_{t(h)} h'') = f(h)(h'') = f(h_0)(h'') = f(h_0)(h''') = f(h')(h''') = (\dashv f(h'))\,(\dashv_{t(h')} h''').$$

We use, in turn, the definitions of truncation, consistency, stationarity of the strategies, consistency, and finally, truncation again. (We make use of the root strategy h_0 as a history which is surely succeeded by both h'' and h'''.) $\square$

Based on the two theorems above, one might expect that a stationary frame of stationary strategies would be consistent. However, this is not necessarily so, as can be seen from the following example. Consider the Markov decision problem in Figure 3. For all $h \succcurlyeq \bar{h}$, let:

$$s_1^{\bar{h}}(h) = \begin{cases} (A, A) & \text{if } \omega(h) = \rho, \\ (C, C) & \text{if } \omega(h) = \sigma, \end{cases} \quad \text{and} \quad s_2^{\bar{h}}(h) = \begin{cases} (B, B) & \text{if } \omega(h) = \rho, \\ (D, D) & \text{if } \omega(h) = \sigma. \end{cases}$$

Now, let us define a frame f, so that:

$$f(h) = \begin{cases} s_1^{\bar{h}} & \text{if } \omega(\bar{h}) = \rho; \\ s_2^{\bar{h}} & \text{if } \omega(\bar{h}) = \sigma. \end{cases}$$
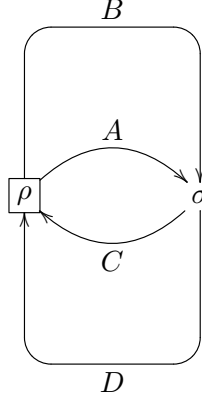
Figure 3: Stationary frame of stationary strategies is not necessarily consistent.

This is obviously a frame of stationary strategies. It also is a stationary frame, since only the end-state matters in assigning a strategy to a history, according to the definition. However, it is not a consistent frame, since:

$$f(\rho)(\rho A\sigma) = s_1^\rho(\rho A\sigma) = (A, A) \neq (B, B) = s_2^{\rho A\sigma}(\rho A\sigma) = f(\rho A\sigma)(\rho A\sigma).$$

3.7 Induced strategy

We assume that, since each agent has control over his current action (and only that), the actual actions executed by each agent h can be obtained from frame f by looking at $f(h)(h)$, in other words, from the diagonal of the frame.

Definition 14. The *induced strategy* of a frame f specifies the actual actions chosen by each agent:

$$\Lambda(f) : h \in H \mapsto A_{\omega(h)}, \text{ given by } \Lambda(f)(h) = f(h)(h).$$

It is handy to define for some frame f, and an agent at $\bar{h}$, the induced strategy for the (present and) future:

$$\Lambda^{\succcurlyeq \bar{h}}(f) : h \in H^{\succcurlyeq \bar{h}} \mapsto A_{\omega(h)}, \text{ given by } \Lambda^{\succcurlyeq \bar{h}}(f)(h) = f(h)(h);$$
$$\Lambda^{\succ \bar{h}}(f) : h \in H^{\succcurlyeq \bar{h}} \setminus \{\bar{h}\} \mapsto A_{\omega(h)}, \text{ given by } \Lambda^{\succ \bar{h}}(f)(h) = f(h)(h).$$

The induced strategy of the frame represented in Table 1 is $\Lambda(f)(\rho) = (BB)$, $\Lambda(f)(\rho A\rho) = (BB)$, $\Lambda(f)(\rho B\rho) = (AA)$ etc.

The language developed in this section can be helpful for discussing most problems of dynamic inconsistency formulated in the multiple-agent framework, and is independent of the specific assumptions on utility functions of the next section. In particular we believe that our (conceptual) distinction between beliefs and intentions, the agent (and its strategy) and the decision maker (and the frame used to describe him), and our method of representing frames

through tables should prove useful. Moreover, the distinction between stationary strategies and stationary frames is, as far as we are aware, entirely new in the literature.

4 Utility and discounting

The term *payoff*, introduced in Definition 1, refers to the immediate gains or losses resulting from an action. Formally, a payoff gained in period t is denoted by u_t, and a stream of payoffs starting at period t by $u_{t\rightarrow} = (u_t, u_{t+1}, \dots)$. We say that a stream of payoff $u_{t\rightarrow}$ starting at period t coincides with a stream of payoff $u'_{t'\rightarrow}$ starting at period t' if $u_t = u'_{t'}$ and $u_{t+1} = u'_{t'+1}$, and so on.

Payoffs are fully determined by the decision problem, the state and the action taken. However, time preference implies that identical payoffs might be regarded differently by various agents. Throughout the paper, we make two assumptions on the utility functions $U^h(u)$ that integrates a stream of future payoffs into a single number. The first assumption says that U^h is continuous for every h.

Assumption 1. U is continuous at infinity – i.e., for any ϵ, there is a horizon T such that the total maximum utility attainable after T is less than ϵ. This is the same assumption as in Fudenberg and Levine (1983).

Another crucial assumption that we use is that agents are identical in the way they evaluate streams of payoffs.

Assumption 2. For any two agents at histories h and h', and coinciding streams of payoffs $u_{t(h)\rightarrow}, u'_{t(h')\rightarrow}$, the utilities of the two agents are equal, i.e., $U^h(u_{t(h)\rightarrow}) = U^{h'}(u'_{t(h')\rightarrow})$.

If the utility functions satisfy First and Second Order separability (Lapied and Renault, 2012), then the discount factor for a future payoff u_t can only depend on the time distance $t(h) - t$. In our examples in Section 7 and the appendix, we use a discounted utility function of a particular form, namely, quasi-hyperbolic discounting:

$$U^h(u_{t(h)\rightarrow}) = u_{t(h)} + \beta \sum_{t=t(h)+1}^{\infty} \delta^{t-t(h)} u_t,$$

with $0 \leq \beta \leq 1$ and $0 \leq \delta < 1$.

In Section 3.3, we distinguished between two senses of the term "expected utility": utility *actually experienced* from expecting a future payoff stream; and "expected utility" as simply *a means of calculating* with various future courses of action. This distinction is formally nailed down and further refined by the following definitions.[13]

---

[13]Sáez-Martí and Weibull (2002) connect discounting with altruism towards future selves. They note that

Definition 15. The *expected utility based on intentions* of playing strategy s^h for an agent at h is:

$$U_i^h \left(s^h \right) = \mathbb{E} \left[i^h \right] \left(U^h \right).$$

On the other hand, whenever an agent is reflecting on how much utility he can reasonably expect, he will calculate his utility based on his beliefs.

Definition 16. The *expected utility based on beliefs* of playing strategy s^h for an agent at h is:

$$U_b^h \left(s^h \right) = \mathbb{E} \left[b^h \right] \left(U^h \right).$$

Neither his intentions, nor his beliefs determine the *real* utility of an agent. When calculating an agent's real (expected) utility, the only thing that matters is which actions future agents will actually implement under various eventualities. The definition of induced strategy captures just this, and can thus be used to define *induced utility*:

Definition 17. Given a frame f, the (ex post) *induced utility* of the root agent at h_0 is:

$$U_r \left(f \right) = \mathbb{E} \left[\Lambda^{\succcurlyeq h_0}(f) \right] \left(U^{h_0} \right).$$

We will use the notion of induced utility for welfare comparisons between various frames, especially in the appendix, when contrasting naive and sophisticated decision makers in particular decision problems. Implicitly, this means picking the perspective of the root agent for welfare comparisons of frames. We do so in order to avoid any normative assumptions on the agent's long-run preferences.[14]

Notice that traditionally, the above three meanings of the term "expected utility" coincide. The reason is that where dynamic inconsistency does not pose a problem, intentions and beliefs on future actions coincide; moreover, the decision maker always executes the intentions of past agents.

5   Naiveté

At first sight, it is not even clear whether naiveté is a property of the decision maker or that of an agent. In this and the following section, we define naiveté and sophistication primarily for agents, assuming that the decision maker is always naive (or sophisticated). We return to the issue presented in Section 1 after analysing and contrasting these base cases.

Naiveté has been characterized in several ways in the literature; a naif is aware or unaware of different things, depending on the particular interpretation. A naif is said to:

"[c]urrent welfare or 'total utility', so defined, does not stem only from current instantaneous utility but also from (the anticipation of) the stream of future instantaneous utilities". Our distinction highlights exactly this: is it the actual future stream, or the anticipation of that stream that is really important?

[14]One alternative would be using a Pareto criterion, as for instance, in Herings and Rohde (2006).

- choose at each stage an option which seems currently the best (Strotz, 1956; Hammond, 1976);

- fail to realize that future selves will have different preferences (O'Donoghue and Rabin, 2001; Sarafidis, 2004; DellaVigna and Malmendier, 2006; Herings and Rohde, 2006; Heidhues and Kőszegi, 2009);

- believe that – though his preferences might change – he has perfect self-control about the future, allowing him to commit to a strategy (O'Donoghue and Rabin, 1999a; Gruber and Kőszegi, 2001; Ali, 2011).

It is easy to see that these are genuinely alternative interpretations of naiveté. A common aspect is that something is wrong with the beliefs held by the agent. We argue that these troubles arise from the way the naif determines his beliefs: particularly, that for a naive agent, his current preferences determine his intentions, which in turn determine his beliefs on future actions. Thus, it does not matter whether the agent holds an explicit belief on the lack of change in his preferences, or whether he believes he will simply fail to act on such changes, or that he has strong beliefs in his own will- or pre-commitment power. The essential features of naiveté are the directions of determination seen in Figure 4. All the above cases are described by this model.

$$\text{Preferences} \rightarrow \text{intentions} \rightarrow \text{beliefs} \rightarrow \text{strategy.}$$

Figure 4: The forming of intentions and beliefs by a naif.

Definition 18. A strategy $\tilde{s}^h$ of an agent at h is *naively optimal*, if it maximizes expected utility based on intentions:

$$\tilde{s}^h \in \arg\max_{s \in S^h} U_i^h(s),$$

and it is coherent:

$$\tilde{b}^h(h') = \tilde{i}^h(h'), \text{ for all } h' \in H^{\succcurlyeq h}.$$

A frame $\tilde{f}$ is naively optimal if the strategy $\tilde{f}(h)$ is naively optimal at each history h.

Naiveté is then primarily a property of an agent. We talk about a naive decision maker if the frame describing him is naively optimal, i.e., if all his agents are naive.

Working in a continuous-time discounted utility framework, Strotz (1956) shows that only when the discount function is exponential does the decision maker possess a consistent naively optimal frame for all decision problems. For any non-exponential discount functions there are decision problems for which there is no consistent naively optimal frame.[15]

---

[15]Blackorby et al. (1973) notes that "among the many unpleasant features of inconsistent planning is that [... the naive's] behaviour, ex post, makes no sense from any point of view."

Theorem 3. *For any decision problem, there exists a stationary naively optimal frame.*

Proof. For each h, the set $\arg\max_{s \in S^h} U_i^h(s)$ is non-empty, because S^h is non-empty and closed for pointwise limits, and U_i^h is continuous (Assumption 1). Therefore, the set of strategies where the maximum is, in fact, reached is non-empty. But note that the optimality condition in the definition of naively optimal strategies only determines the intention-component of strategies, thus, beliefs can be constructed freely. This means that we can ensure coherence, i.e., we can choose a naively optimal strategy at each h.

Now, to guarantee that the generated frame is stationary, we need to choose the same truncated strategy for each set of histories where the end-state is identical. This is always possible, since whenever the final state is identical for two histories, both the truncated strategy set and the utility function defined at those histories are identical (per our assumption on utility functions), and therefore so are the set of truncated optimal strategies. $\square$

Is the stationary naively optimal frame unique? For a naive decision maker, this depends on whether for each history, there is a unique naively optimal strategy; this latter problem can be reduced to whether for each *state* that can be reached, there is a unique naively optimal strategy (defined at any history where the current state is that state). For generic decision problems, it seems likely that this is indeed the case, though the scope of this proof is beyond this paper. Now, if there is a unique naively optimal strategy for each history, then there is only one naively optimal frame – and it is stationary, too. In degenerate cases, where multiple naively optimal strategies can be assigned to at least one state, we get stationary naively optimal frames, along with non-stationary ones. It should also be noted that – because of the possibility of inconsistency – multiplicity of naively optimal frames also leads to a multiplicity of induced utilities. In particular, the agents facing multiple naively optimal strategies believe it does not matter which strategy they choose, as they expect they will stick to those strategies – but they can be wrong.

As an example, consider the decision problem on Figure 5. Discounting is quasi-hyperbolic with $\beta = \delta = 0.5$. Since naively optimal strategies are coherent, we will only consider such strategies. There are only two histories, $h_0 = (\rho)$ and $h_1 = (\rho, A, \sigma)$, where the action choice is not trivial, so the naive root agent at $h_0 = (\rho)$ has to consider only four coherent strategies, which we will denote, by $s_{AC}^{h_0}, s_{AD}^{h_0}, s_{BC}^{h_0}$ and $s_{BD}^{h_0}$. They are defined through:

1. $s_{AC}^{h_0}(h_0) = (A, A)$ and $s_{AC}^{h_0}(h_1) = (C, C)$;

2. $s_{AD}^{h_0}(h_0) = (A, A)$ and $s_{AD}^{h_0}(h_1) = (D, D)$;

3. $s_{BC}^{h_0}(h_0) = (B, B)$ and $s_{BC}^{h_0}(h_1) = (C, C)$;

4. $s_{BD}^{h_0}(h_0) = (B, B)$ and $s_{BD}^{h_0}(h_1) = (D, D)$.

It can be easily seen that from the perspective of the root agent, $U_i^{h_0}(s_{AC}^{h_0}) = U_i^{h_0}(s_{BC}^{h_0}) = U_i^{h_0}(s_{BD}^{h_0}) = 0 > -\frac{1}{4} = U_i^{h_0}(s_{AD}^{h_0})$. Thus, the naive root agent is indifferent between choosing action A first, and C afterwards, or simply B.
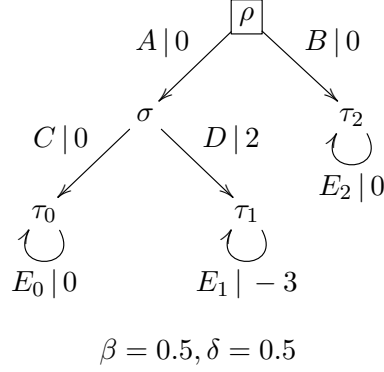

$$\beta = 0.5, \delta = 0.5$$

Figure 5: Multiplicity of induced utility for naively optimal frames.

Now we focus on the agent at h_1. He has two strategies available, which we will denote by $s_C^{h_1}$ and $s_D^{h_1}$. They are defined through:

1. $s_C^{h_1}(h_1) = C$;

2. $s_D^{h_1}(h_1) = D$.

For the agent at h_1, $U_i^{h_1}(s_D^{h_1}) = 0.5 > 0 = U_i^{h_1}(s_C^{h_1})$. Therefore, the naive agent at h_1 prefers to choose action D. Now, consider the two naively optimal frames f_{BCD} and f_{ACD}, defined through:

- $f_{BCD}(h_0) = s_{BC}^{h_0}$ and $f_{BCD}(h_1) = s_D^{h_1}$;

- $f_{ACD}(h_0) = s_{AC}^{h_0}$ and $f_{ACD}(h_1) = s_D^{h_1}$;

From the previous calculations, we see that both f_{BCD} and f_{ACD} are naively optimal frames. However, the induced utilities are not equal: $U_r(f_{BCD}) = 0$ and $U_r(f_{ACD}) = -\frac{1}{4}$. The underlying reason is the following: when an agent picks a naively optimal strategy in h_0, he expects to earn 0 by either choosing strategy $s_{BC}^{h_0}$ or $s_{AC}^{h_0}$. But if he chooses $s_{AC}^{h_0}$, after getting to history h_1 – on account of being present-biased – he will not stick to his previous strategy, which would prescribe him choosing action C; instead, he chooses D, which leads to a decrease in his induced utility.

To sum up our discussion of naiveté: if we interpret an optimal frame as predictive for a naive decision maker's behavior, then in the generic case, we get a unique prediction of actions, intentions and belief for each history; and we expect stationary behavior in realization, and

a single prediction for the induced utility. On the other hand, in degenerate cases we can get multiple predictions of actions, intentions and beliefs for some agents; we do not necessarily expect stationary behavior; and we do not necessarily get a unique expectation for induced utility.

6   Sophistication

Similar to naiveté, there are several definitions of sophistication:

- optimality under a credibility constraint: following a feasible optimal strategy, or a plan that he will actually follow (Strotz, 1956; Yaari, 1977);

- game-theoretic notion: an intra-personal subgame-perfect equilibrium, sometimes also referred to as Strotz-Pollak equilibrium (Kocherlakota, 1996; Vieille and Weibull, 2009);

- rational expectations: perfectly anticipating future behavior (O'Donoghue and Rabin, 2001; Gilpatric, 2003);

- self-awareness: being aware of future changes in discount rates and preferences (Hammond, 1976; Heidhues and Kőszegi, 2009; Ali, 2011).

These definitions do not match precisely. For instance, the notion of sophistication as an intra-personal equilibrium does not guarantee the satisfaction of rational expectations in case when multiple such equilibria exist. Just like for naiveté, we would like to offer a new interpretation of sophistication. We regard sophistication as primarily a property of agents. In our framework, the defining characteristic of sophisticated agents is that they first consider their beliefs about the future, and only then do they form intentions. For now, our sophisticated agents assume that all future agents will be sophisticated, too – we will relax this assumption for hybrid agents in Section 7.

A sophisticatedly optimal strategy is made up, first, by beliefs about future agents' choices: he believes each future agent will pick a best response to future agent's choices. So we implicitly have to consider second-order beliefs, the beliefs of each agent about the beliefs of agents about the future. However, we assume that in a sophisticated strategy, the second-order beliefs coincide with first-order beliefs.[16] Thus, the current agent believes that future agents will believe what he currently believes.

The second component of a sophisticatedly optimal strategy concerns the intentions, which are set to match the beliefs; as the agent knows he has no control over his future selves, he can just as well intend future actions that they are choosing anyway.

---

[16]O'Donoghue and Rabin (2001) makes the assumption on second-order beliefs explicit for partially naive agents.

Beliefs and preferences $\rightarrow$ intentions $\rightarrow$ action $\rightarrow$ strategy.

Figure 6: The forming of intentions and beliefs by a sophisticated.

Finally, the current action is chosen to be a best response to future actions.

Definition 19. A strategy $\hat{s}^h$ of an agent at h is *sophisticatedly optimal*, if:

$$\hat{b}^h\left(h'\right) \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}\left(\hat{s}^h\left[a : h'\right]\right), \text{ for all } h' \in H^{\succcurlyeq h},$$

and it is coherent:

$$\hat{i}^h\left(h'\right) = \hat{b}^h\left(h'\right), \text{ for all } h' \in H^{\succcurlyeq h},$$

where $\hat{s}^h\left[a : h'\right]$ denotes the strategy where the action taken at h' is replaced with action a in strategy $\hat{s}^h$.[17]

A frame $\hat{f}$ is sophisticatedly optimal if the strategy $\hat{f}(h)$ is sophisticatedly optimal at each history h.

One remark about the intention component of a sophisticatedly optimal strategy is in order. It could be argued that instead of intending to give a best response to future beliefs at h' based on the preferences at h', the agent at h should intend something else at h'; namely, to give a best response to the choices of future agents based on the preferences at h, not the ones at h'. However, we want to ensure coherence of sophisticatedly optimal strategies to ensure that a consistent sophisticated strategy always exists. Moreover, it is psychologically plausible that the sophisticated agent wants to maintain this coherence – indeed, this is why he is reasoning about future agents after all, realizing that the preferences of future agents will be different from his current ones. We will see in Section 7 that unlike for purely naive or sophisticated types, for hybrid agents intentions and beliefs might not match.

Theorem 4. *For any decision problem, there exists a consistent sophisticatedly optimal frame.*

Proof. Fix an enumeration of all histories, that is, a bijection $\rho : \mathbb{N} \to H$. Although it is not necessary for the proof, we can assume that the root history is taken first, then all histories at stage 1 are enumerated, then all histories at stage 2, and so on. Let $\mathscr{A} = \times_{h \in H} A_{\omega(h)}$, where the product is taken in the order according to ρ. So, an element $a = \left(a^h\right)_{h \in H}$ of $\mathscr{A}$ prescribes, for every history $h \in H$, an action a^h for the agent at h. Take an arbitrary $\hat{a} \in \mathscr{A}$.

Now, for every $n \in \mathbb{N}$, we construct an $a_n \in \mathscr{A}$ as follows. For every history h beyond stage n, that is, with $t(h) > n$, let $a_n^h = \hat{a}^h$. Then, we proceed by backwards induction. For every

---

[17]When calculating $U_b^{h'}\left(s^h\right)$, only the payoffs generated by s^h for the histories succeeding h' should be taken into account; i.e., we consider the expected utility induced by s^h for the subtree starting at h'.

history h at stage n, that is $t(h) = n$, let a_n^h be an action for the agent at h that maximizes his utility if all agents that succeed him play the action according to a_n (or equivalently, according to $\hat{a}$). In general for a $k \in \{1, \dots, n\}$, if we have defined a_n^h for all histories h with $t(h) > k$, then for every history h with $t(h) = k$, we choose a_n^h to be an action for the agent at h that maximizes his utility if all agents that succeed him play the action according to a_n.

So, we obtain a sequence $(a_n)_{n \in \mathbb{N}}$ in the space $\mathscr{A}$. Note that $\mathscr{A}$ with the product topology is compact, and because H is countable, it is metrizable too. Consequently, it is sequentially compact, which implies that the sequence $(a_n)_{n \in \mathbb{N}}$ has a subsequence $(a_{n_k})_{k \in \mathbb{N}}$ which converges to some $\hat{a} \in \mathscr{A}$. This means that, for every $m \in \mathbb{N}$, there exists an $K_m \in \mathbb{N}$ such that, for every $k \geq K_m$, the actions $a_{n_k}^h$ and $\hat{a}^h$ coincide for all histories h with $t(h) \leq m$.

Let s_n^h be the strategy of the agent at history h which intends to play the action $a_n^{h'}$ at all histories $h' \succeq h$ and believes that these actions will be played – thus, a coherent strategy. Similarly, we define the strategy $\hat{s}^h$ with respect to $\hat{a}$. Now consider the frame $\hat{f}$ that assigns strategy $\hat{s}^h$ to the agent at h, for every history h.

First, we show that $\hat{f}$ is consistent. Take any h, h', the corresponding strategies $\hat{s}^h$ and $\hat{s}^{h'}$ and some history h'' with $h'' \succcurlyeq h$ and $h'' \succcurlyeq h'$. Then

$$\hat{s}^h \left(h'' \right) = \left(\hat{a}^{h''}, \hat{a}^{h''} \right) = \hat{s}^{h'} \left(h'' \right).$$

Thus, $\hat{f}$ is a consistent frame.

We now prove that $\hat{f}$ is sophisticatedly optimal. For this purpose, consider an arbitrary agent, say at history h, and an agent at a history $h' \succeq h$. By construction, for every $n \geq t(h')$, the action $a_n^{h'}$ maximizes the utility of the agent at h' if all agents that succeed him play the action according to a_n. Thus,

$$b_n^h \left(h' \right) = a_n^{h'} \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'} \left(s_n^h \left[a : h' \right] \right)$$

and as we have seen, s_n^h is coherent:

$$i_n^h \left(h' \right) = b_n^h \left(h' \right).$$

By taking the limit along the subsequence $(n_k)_{k \in \mathbb{N}}$ and using continuity, we obtain

$$\hat{b}^h \left(h' \right) = \hat{a}^{h'} \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'} \left(\hat{s}^h \left[a : h' \right] \right)$$

and

$$\hat{i}^h \left(h' \right) = \hat{b}^h \left(h' \right).$$

Thus, $\hat{f}$ is a consistent sophisticatedly optimal frame indeed. $\qquad\square$

Again, we might ask whether there is a unique consistent sophisticatedly optimal frame. In contrast with naively optimal frames, it turns out that this is not the case, even for generic decision problems. In the problem of underinvestment (Section A.2, case 1.b.ii) we get multiple sophisticatedly optimal strategies for each history. Moreover, the same example shows that there might be no stationary sophisticatedly optimal frame when we only allow for pure actions. Also, the induced utilities of optimal frames can differ. Thus, only in some cases do we get a unique sophisticatedly optimal frame, and with it, unique predictions for a sophisticated decision maker's actions, intention, beliefs and induced utility.[18] Luckily, for these cases Theorem 4 implies that the frame will be consistent, too.

Overall, our understanding of sophistication in terms of sophisticated agents first working through their beliefs, then deriving their intentions is closest to the self-awareness interpretation. However, sophisticated optimality can indeed be regarded as a notion of intra-personal subgame-perfection.[19] As there might be several such equilibria, agents at various histories might pick actions corresponding to different equilibria; thus, there is no a priori guarantee for the satisfaction of rational expectations, or that the subgame-perfect equilibrium chosen by the root agent will actually be followed through. Thus, the equilibrium aspect of sophisticated optimality on the *strategy* level does not imply consistency or stationarity for the sophisticated *frame*. However, the above theorem shows that a consistent sophisticatedly optimal frame exists.

The next natural question concerns the relative advantages of sophistication against naiveté.[20] The most commonly analyzed examples in the literature of dynamic inconsistency are stories of procrastination, impulsiveness, underinvestment, addiction and binging behavior. We can classify these problems into two groups: in the first group, there are decisions that concern the execution of a single task, like finishing an academic paper or ending a marriage in a sudden burst of anger (procrastination and impulsiveness). In the other conceptual box, we put problems concerning the *repeated* execution of a task, like not saving enough for retirement, or drug addiction (underinvestment and binging). It turns out that we can model all of these as Markov decision problems with only two states (Figures 9 and 10), if we assume a decision maker that uses quasi-hyperbolic discounting, arguably the simplest form of non-exponential discounting.

---

[18]On the non-uniqueness of sophisticatedly optimal strategies, see Phelps and Pollak (1968); Peleg and Yaari (1973); Blackorby et al. (1973). More recently Vieille and Weibull (2009) show that non-uniqueness is a generic property for quasi-hyperbolic discounting, and also give sufficient conditions for uniqueness. For a refinement concept, see Kocherlakota (1996).

[19]This also highlights why in Definition 5 we defined a strategy as an intention-belief pair for *all* future histories, since – as in subgame-perfection – the actions of agents at histories off the optimal path are relevant.

[20]For a the detailed comparison and the calculations, we refer the reader to the appendix. Here we summarize the main findings of the analysis contained therein.

In the appendix, we derive conditions for the payoffs, the discount factor, and the present-biasedness parameter to classify these problems. We find that in some decision problems involving either a single or a repeated task, namely, impulsivity and binging problems, sophisticates commit the same mistake as naifs, and sophistication brings no benefits. This should not come as a surprise, as we cannot expect a notion of intra-personal subgame-perfect equilibrium to solve all welfare problems of decision makers. In some single and also in some repeated decision problems, however sophistication brings clear benefits, and from the perspective of the root agent, the induced utility of a sophisticatedly optimal frame can be strictly higher than that of naively optimal one. Again, this is something to be expected – sophistication would hardly deserve the attention it receives if it never brought any improvement over naiveté.

Coincidentally, among our examples, decision problems in which sophistication is beneficial are also the ones where there exists no sophisticated frame of stationary strategies. Thus, these cases serve as examples to the point that a sophisticated agents might be forced to use non-stationary strategies when mixed actions are disallowed. For the underinvestment problem, we also analyze the possibility of using mixed actions, and calculate the mixing probability that leads to a consistent sophisticated frame in stationary strategies.

The final section of the appendix presents the indulgence problem, discovered by O'Donoghue and Rabin (1999a). This is the simplest known decision problem where naiveté outperforms sophistication from the perspective of the root agent. Thus, the advantages of sophistication are shown not be equivocal across all decision problems, and changing to sophistication might harm the decision maker. We make use of this finding in constructing our final example of hybridity decision making in the next section.

7 Hybrid decision makers

This section introduces a new type of decision maker. So far, a naive decision maker is always naive, a sophisticated is always sophisticated. However, such purity is quite rare, perhaps even nonexistent in the real world. Even the most naive individual realizes after a while that his intentions might not be credible; and even the most consistently sophisticated individual can slip into wishful thinking about his future actions. Therefore, we try to model this duality of an individual through hybrid types. In contrast to the partially naifs of O'Donoghue and Rabin (2001) who (in the context of quasi-hyperbolic discounting) are aware of future present-biasedness, but underestimate its magnitude, our hybrid decision maker flip-flops between being sophisticated and naive, like Ulysses before arriving to the sirens and while listening to them.

We now extend our model to capture hybrid types. Our type space includes naifs and sophisticates.

Definition 20. A *Markov decision problem with agent types* is made up of:

- the set of time periods $\{0, 1, 2, \dots\}$;

- $\Theta \subseteq \Omega \times X$, where Ω is a finite state space and $X = \{N, S\}$ is the finite type space[21]; we denote a state-type pair by θ; the initial state-type pair is $\bar{\theta} \in \Theta$; the state component is denoted by $\omega(\theta)$; while the type component is denoted by $x(\theta)$;

- a finite and non-empty set of pure actions A_ω that the decision maker can choose from in ω;

- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action in state ω;

- transition probabilities $m_\theta : A_{\omega(\theta)} \to \Delta(\Theta)$, with $m_\theta\left(\theta' | a_{\omega(\theta)}\right)$ denoting the probability to transit from the state-type pair θ to the state-type pair θ' when action $a_{\omega(\theta)}$ is chosen.

This definition keeps the Markovian properties of the original model, and adds a specification of naiveté or sophistication to each state. Also, Θ is common knowledge among the agents.

Our model is quite general. Before moving on to illustrate its use in detail for our motivating example from Section 1, we list a few kinds of decision problems for which it could be used:

- exogeneous types: all state-type combinations are allowed ($\Theta = \Omega \times X$). Moreover, in each new state, the agent is naive (or sophisticated) with the same probability: $m_\theta((\omega', x)|a_{\omega(\theta)}) = m_\theta((\omega'', x)|a'_{\omega(\theta)})$ for all $\theta, \omega, \omega', x$. Such a model requires there being no correlation between state and type.

- fixed type for each state: here, we have $(\omega, x), (\omega, x') \in \Theta \Rightarrow x = x'$. This is the opposite of the previous scenario, as there is perfect correlation between state and type. It can be easily seen that any Markov decision problem with agent types can be reformulated in this manner by expanding the state space; however, it can make the model less illuminating, possibly hiding structural similarities between the problems faced by a naive and a sophisticated agent.

- deterministic type determination: this requires that for all $\theta, a_{\omega(\theta)}$, there is some $x \in X$, such that $\sum_{\omega'} m_\theta((\omega', x)|a_{\omega(\theta)}) = 1$. This means that whatever the agent chooses, his type (but not necessarily his state) in the next period is fully determined.

---

[21] N stands for naive, and S for sophisticated.

- full control over type: for all $\theta, x \in X$, there is some $a_{\omega(\theta)}$ so that $\sum_{\omega'} m_\theta((\omega', x)|a_{\omega(\theta)}) = 1$. Each agent can always ensure his type to be whatever he wants for the next period.

Of course, these are merely edge cases, and many interesting situations lie in the middle, having some, but imperfect correlation between state and type; and giving some, but less than total control for agents over their future types. Indeed, it could be argued that the drinking problem we present below should involve stochastic, instead of deterministic type determination. However, for simplicity of analysis, we abstract from such complications.

To understand optimal strategies for hybrid agents, we first have to re-define the notion of a history:

Definition 21. A *type-dependent history* h has the form $h = \big(\theta_0, a_{\omega(\theta_0)}, \ldots, \theta_{t-1}, a_{\omega(\theta_{t-1})}, \theta_t\big)$, with:

- $\theta_i \in \Theta$, for $i \in \{0, 1, \ldots, t\}$, and $\theta_0 = \bar{\theta}$;

- $a_{\omega(\theta_i)} \in A_{\omega_i}$, for $i \in \{0, 1, \ldots, t - 1\}$;

- $m_{\theta_i}\big(\theta_{i+1}|a_{\omega(\theta_i)}\big) > 0$, for $i \in \{0, 1, \ldots, t - 1\}$.

Extending the previous notation, $x(h)$ refers to the current type. We keep the association between histories and agents – each history now corresponds to an agent, and it also includes the agent's type.

Next, we define optimal type-dependent strategies for the Markov decision problem with agent types. We make use of Definitions 18 and 19 for naively and sophisticatedly optimal strategies. We first present the formal definition, and then explain the intuitions below.

Definition 22. A type-dependent strategy $\breve{s}^h$ for a Markov decision problem with agent types is *optimal* at history h, if it satisfies the following conditions:

- for $x(h) = N$:
$$\breve{s}^h \in \arg\max_{s \in S^h} U_i^h(s),$$
and
$$\breve{b}^h(h') = \breve{i}^h(h'), \text{ for all } h' \in H^{\succcurlyeq h}.$$

- for $x(h) = S$:
$$\breve{b}^h(h') \in [\arg\max_{s \in S^{h'}} U_i^{h'}(s)](h'), \text{ for all } h' \in H^{\succcurlyeq h} \text{ with } x(h') = N;$$

$$\breve{b}^h(h') \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}\big(\breve{s}^{h'}[a : h']\big), \text{ for all } h' \in H^{\succcurlyeq h} \text{ with } x(h') = S;$$

and

$$\check{i}^h \left(h'\right) \in \underset{a \in A_{\omega(h')}}{\arg\max} \, U_b^{h'} \left(\check{s}^{h'} \left[a : h'\right]\right), \text{ for all } h' \in H^{\succcurlyeq h} \text{ with } x\left(h'\right) = N;$$

$$\check{i}^h \left(h'\right) = \check{b}^h \left(h'\right), \text{ for all } h' \in H^{\succcurlyeq h} \text{ with } x\left(h'\right) = S.$$

A type-dependent frame $\check{f}$ is optimal, if $\check{f}(h)$ is an optimal type-dependent strategy for all h.

Although this definition is rather lengthy, it captures our basic intuitions for the two types. A naive agent at h does not reason about future agents, as his intentions determine his beliefs, and thus, his whole optimal strategy in the standard way. However, a sophisticated agent at h is able to reason about future agents in the following manner: if a future agent at h' is naive, then the agent at h believes the agent at h' will act in a naive way, maximizing his expected utility based on intentions. If, on the other hand, a future agent at h' is sophisticated, then the agent at h (correctly) believes that the agent at h' will act in a sophisticated way, being able to reason about future agents just as well as h himself does. So a sophisticated agent at h intends to choose in a sophisticated manner at all nodes, giving a best response to the choices of future agents. This implies that the intention and belief component of an optimal type-dependent strategy match for all future histories where the agent is sophisticated, but they might not match for future histories where the agent is naive. Coherence is therefore not a property of optimal type-dependent strategies.

Since naively optimal strategies are, in general, not made up of stationary strategies, the belief-component of an optimal type-dependent frame can also be non-stationary. Moreover, it is easy to see that such a frame is not consistent: the intentions of sophisticated agents will not, in general, correspond to the actions taken by naive agents. In both examples of hybrid decision making we present below, we will see such inconsistencies.[22] The lack of coherence, stationarity and consistency represent the inner conflicts that arise within a hybrid decision maker.

Theorem 5. *For any decision problem, there exists a type-dependent optimal frame.*

Proof. First, start with determining the intentions and beliefs of naive agents, i.e., those at histories where $x(h) = N$. For these, we can simply use the first part of the proof of Theorem 3. So, we have $\check{s}^h = (\check{i}^h, \check{b}^h)$ defined for all h with $x(h) = N$. Let $\check{f}(h) = \check{s}^h$ for all such h.

Moving now to histories where the agent is sophisticated, our construction is analogous to that of Theorem 4. Let $\mathscr{A} = \times_{h \in H} A_{\omega(h)}$, and take an arbitrary $a = (a^h)_{h \in H} \in \mathscr{A}$ that assigns an action to each history.

---

[22]Recall that when an action is actually taken by an agent, it is both believed and intended by the agent for the current history.

Recall that a type-dependent optimal frame will not necessarily be coherent. Therefore, both intentions and beliefs have to be constructed; we start with beliefs. First, transform a into $\hat{a}$ by fixing actions assigned to histories where the agent is naive, i.e., where $x(h) = N$, so that they are actions corresponding to those agents acting in a naive way:

$$\hat{a}^h = \breve{i}^h(h) \quad \text{if } x(h) = N;$$
$$\hat{a}^h = a^h \qquad \text{if } x(h) = S.$$

Let $a_n \in \mathscr{A}$ be defined as follows: $a_n^h = \hat{a}^h$ for all h with $t(h) > n$, or with $t(h) \leq n$, and $x(h) = N$. For the remaining histories with $t(h) \leq N$ and $x(h) = S$, we move by backwards induction from n to 0, and let a_n^h be the best response to the future actions, which are already all defined.

Thus, we obtain a sequence $(a_n)_{n \in \mathbb{N}}$ in $\mathscr{A}$. Since $\mathscr{A}$ is sequentially compact (see the proof of Theorem 4), the sequence a_n has a subsequence a_{n_k} converging to some $\hat{a} \in \mathscr{A}$.[23] Thus, for all possible choices of a horizon m, we can find a K_m such that $\hat{a}^h = a_{n_k}^h$ for all $k \geq K_m$ and h with $t(h) \leq m$.

Now, let the beliefs of a sophisticated agent – i.e., at a history h with $x(h) = S$ – be $\breve{b}^h(h') = \hat{a}^{h'}$.

Finally, we construct the intentions of sophisticated agents. Set sophisticated agents' beliefs about future nodes as: $\breve{i}^h(h') = \breve{b}^h(h')$ for all h' with $x(h) = x(h')$. For sophisticated agents' intentions assigned to future naive nodes, we construct $\bar{a}$ by modifying $\hat{a}$ in a way that whenever $x(h) = N$, $\bar{a}^h$ is a best response to the future actions in $\hat{a}$. We keep actions at other, sophisticated histories unchanged: $\bar{a} = \hat{a}$. We set the intentions of a sophisticated agent to be $\breve{i}^h(h') = \bar{a}^{h'}$.

We have thus defined both $\breve{b}^h$ and $\breve{i}^h$ for sophisticated agents. Let $\breve{s}^h = (\breve{i}^h, \breve{b}^h)$ and finally, $\breve{f}(h) = \breve{s}^h$ for all $x(h) = S$. This completes our construction of a type-dependent optimal frame, as we have provided the strategies assigned to histories where the agent is naive, and also to the ones where he is sophisticated.

We will now review our construction again to confirm that $\breve{f}$ is indeed a type-dependent optimal frame. For $x(h) = N$, this is immediate. For $x(h) = S$, we will first check the beliefs, and then the intentions.

First, suppose that $x(h') = N$. We have:

$$\breve{b}^h(h') = \hat{a}^{h'} = \breve{i}^{h'}(h') \in \arg\max_{s \in S^{h'}} U_i^{h'}(s)(h'),$$

which is what is required in the definition of $\breve{f}$. For the other case, take $x(h') = S$. Now,

[23] There might be multiple subsequences converging to different $\hat{a}$-s; in that case, we can select any one of them.

from our construction of $a_n^{h'}$, for all $n \geq t(h')$:

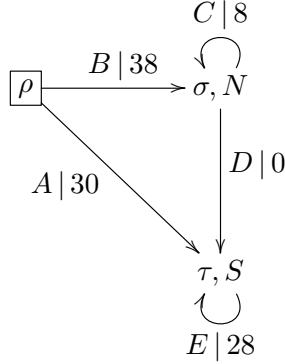$$\check{b}_n^h(h') = a_n^{h'} \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}(b_n^h[a:h']).$$

Taking the limit along the subsequence n_k, using continuity, we get:

$$\check{b}^h(h') = a^{h'} \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}(b^h[a:h']).$$

For the intentions of sophisticated agents for future sophisticated nodes, we set these directly to be $\check{i}^h(h') = \check{b}^h(h')$. The last thing we need to check is the intentions of sophisticated agents for future naive agents. These were defined as:

$$\check{i}^h(h') = \bar{a}^{h'} \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}(\check{b}^{h'}[a:h']).$$

$\square$

$$\beta = 0.5, \delta = 0.5$$

Figure 7: A drinking problem.

We now return to the problem raised in the introduction, displayed in Figure 7. The decision maker is sitting in the pub, having finished his second beer (in state ρ), and is of type S (sophisticated). He can either go home directly, by choosing action A, transiting to state τ, where he does not need to make any more decisions. Alternatively, he can drink "one more beer" by choosing action B. This, however, transitions him to a "drunken" state σ, where he becomes type N (naive). In this drunken state, he can choose between drinking one more beer by choosing C (thus, maintaining his drunkenness and staying at σ), or going home to state τ by choosing D.

Let $h = (\rho)$ be the root agent, and let $h_{drink} = (\rho, B, \sigma)$ be the agent who drank "one more beer". Our goal is to construct $\check{s}^h$, the optimal type-dependent strategy for the root

agent. Since $x(h) = S$, we have to deal with the interesting case, that of a sophisticated root agent.

We start the analysis of this situation by focusing on the state σ. Take any h' such that $\omega(h') = \sigma$. As $x(h') = N$ if $\omega(h') = \sigma$, we get $\breve{b}^h(h') \in \arg\max_{s \in S^{h'}} U_i^{h'}(s)(h')$ according to the definition; the root agent believes that an agent at history h' with a current state σ is naive. What is the naive choice in state σ? It is easy to see[24] that the (subgame-optimal) naively optimal strategy is $(C,C)(D,D)(D,D)\ldots$, i.e., drink one more beer, and then go home. Thus, $\breve{b}^h(h') = C$ whenever $\omega(h') = \sigma$. The root agent thus believes he would continue drinking after becoming naive.

What about the intentions of the root agent for the agent at σ? He believes that at history h', the continuation actions will be C. The agent at history h' could only choose between C and D. Going for C yields $8 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{8}{1 - \frac{1}{2}} = 12$, whereas picking D gives $0 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{28}{1 - \frac{1}{2}} = 14$. Therefore, $\breve{i}^h(h') = D$ whenever $\omega(h') = \sigma$. Together with the result of the previous paragraph, we get that $\breve{s}^h(h') = (D, C)$ whenever $\omega(h') = \sigma$. We see that the intentions and the beliefs of the root agent do not match: he would like future agents to pick D at every h', but correctly anticipates that future agents will be unable to do so, and would actually choose C. The sober, sophisticated root agent realizes that if he drinks just one more beer, he will end up drinking a lot more than he actually wishes for.

So what should the root agent choose at h? He can pick A, going home directly, earning him $U_b^{h_0}(\breve{s}^{h_0}[A:h_0]) = 30 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{28}{1 - \frac{1}{2}} = 44$. Or, he can pick B, drink one more beer, and end up in the pub. This would earn him $U_b^{h_0}(\breve{s}^{h_0}[B:h_0]) = 38 + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{8}{1 - \frac{1}{2}} = 42$. Going home seems best. Thus, $\breve{s}^h(h) = (A, A)$. The optimal type-dependent strategy is thus:

$$\breve{s}^h(h) = (A, A),$$
$$\breve{s}^h(h') = (D, C), \text{ whenever } \omega(h') = \sigma,$$
$$\breve{s}^h(h') = (E, E), \text{ whenever } \omega(h') = \tau.$$

Note that a fully naive root agent would expect that he can resist the temptation of additional beers, and would expect a utility of $38 + \frac{1}{2} \cdot \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2}^2 \cdot \frac{28}{1 - \frac{1}{2}} = 45$. The sophisticated root agent realizes that this is unachievable, as the incentives *and the type* of the agents changes by transiting to σ. So in the drinking problem, the sober, sophisticated root agent avoids becoming naive, and thus is better off. Sophistication thus can help avoiding the trap of naiveté. But can sophistication help in avoiding the pitfalls of sophistication?

Section A.3 shows that in some decision problems – in particular, in the indulgence prob-

---

[24]Obviously, the decision problem reduced to the states σ and τ is that of a procrastination of a single task, discussed in Section A.1, case 1.b. Indeed, with these payoffs, $8 < (1 - \frac{1}{2}) \cdot 0 + \frac{1}{2} \cdot 28 = 14$, and $8 > (1 - \frac{1}{2} \cdot \frac{1}{2}) \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} \cdot 28 = 7$. So a non-present biased agent would go home, but a present-biased agent prefers to postpone it.

lem – a naive decision maker is strictly better off than a sophisticated one. This allows us to construct the decision problem in Figure 8, which we call the indulgence problem with hybrid type.
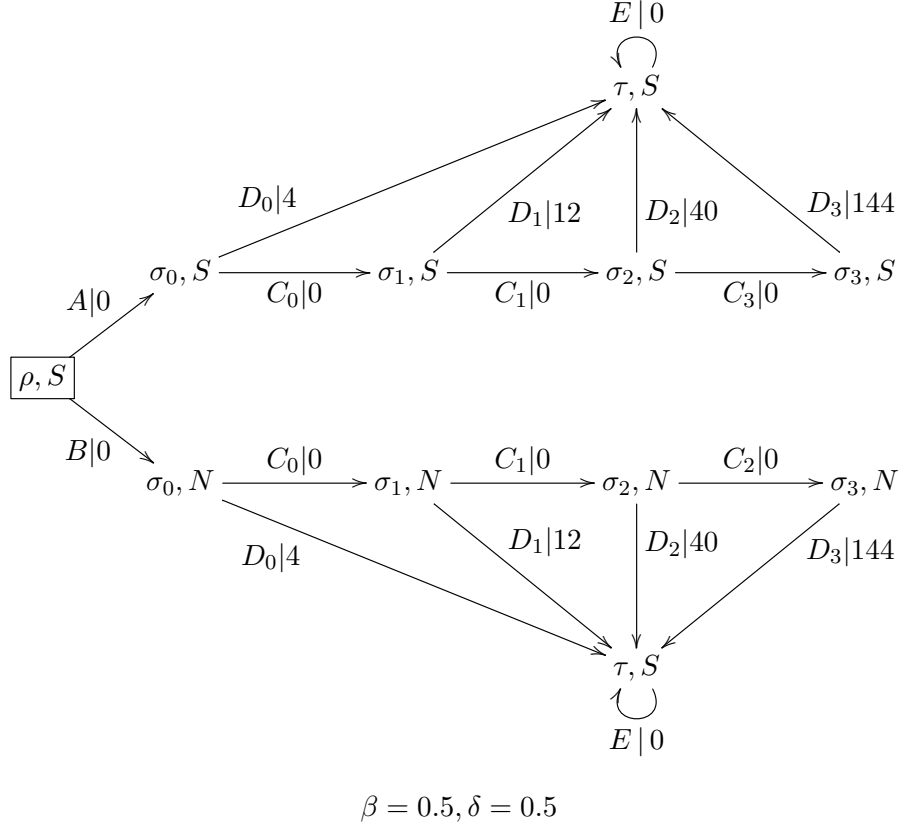

$$\beta = 0.5, \delta = 0.5$$

Figure 8: Optimal self-deception in the indulgence problem with hybrid type.

The agent at the root history $h_0 = ((\rho, S))$ is sophisticated. Therefore, he is able to reason about future agents in the following way: at history $h_1 = ((\rho, S), A, (\sigma_0, S))$, and at all succeeding histories, he will be sophisticated. Thus, he concludes – after an analysis similar to the one in Section A.3 – that he will choose the action D_0. On the other hand, at history $h_1' = ((\rho, S), B, (\sigma_0, N))$, and at all succeeding histories, he will be naive. Thus, he concludes that in the subgame starting at h_1', he will choose action D_2. From the perspective of the root agent:

$$U_b^{h_0}(\check{s}^{h_0}[A : h_0]) = \tfrac{1}{2} \cdot \tfrac{1}{2} \cdot 4 = 1 < 2.5 = \tfrac{1}{2} \cdot \left(\tfrac{1}{2}\right)^3 \cdot 40 = U_b^{h_0}(\check{s}^{h_0}[B : h_0]).$$

Therefore, the best response of the root agent is to choose action B. Notice that this means that a sophisticated agent chooses to face the indulgence problem as a naif, thereby intentionally causing the agent at h_1' to have wrong beliefs. In particular, the naive agent at h_1' believes that he will be able to wait until the wine fully matures, and take action D_3. The

sophisticated agent at the preceeding history h_0 knows that this is not the case, that in fact, action D_2 will be taken. Thus, when choosing an optimal type-dependent strategy, the root agent realizes that he is better off with false beliefs, and decides to deceive himself.

By what means such self-deception might be effectively achieved, or whether it can be achieved intentionally at all is, of course, a difficult problem. But it seems like self-deception has its virtues, which might, in itself, challenge ethical arguments on the inherent immorality of self-deception.[25]

8 Concluding remarks and future research

This paper attempts to play a foundational role for future discourse in multi-self models of dynamic inconsistency. It attempts to establish that the basic epistemic concepts to be considered are beliefs and intentions, and the main levels of analysis should be those of strategies and frames. We would now like to provide some remarks and outline some directions for follow-up research in this area.

An obvious limitation of the current framework is that it only allows for pure actions. This limitation is introduced to ease the presentation, but the technical adaptations required for dealing with mixed actions can be accomplished rather straightforwardly. Mixed actions should play a particularly important role when moving from decision-theoretic models to a game setting.

One might wonder how flexible this model is with regards to increasing the state or action space of the decision problem. Countably infinite states and actions can be allowed for without much difficulty, as long as the set of payoffs for each action remains compact (and hence, bounded). However, handling continuous time would require a fundamentally different framework, along with a reinterpretation of the notion of an "agent".

Whereas our focus was the two most common types of decision makers facing dynamic inconsistency, naifs and sophisticates, there have been arguments in the literature for taking seriously other types as well. In particular, McClennen (1990) argues for the possibility of resolute decision making. In fact, resoluteness can easily be incorporated into our framework. It would be interesting to expand the type space in Markov decision problems with hybrid types to include resolutes.

The horizon of sophisticated decision makers requires further investigation. If agents have only a finite horizon, reasoning about future agents can be based on two assumptions: either the *length*, or the *endpoint* of the horizon of that future agent is the same as that of the current agent. In the former case, we are talking about a *moving*, in the latter, about a *fixed*

---

[25]For an overview on the philosophical problems of self-deception, see Deweese-Boyd (2012).

horizon. The implications of these two assumptions on the optimal strategies (generated, for instance, through a backward induction reasoning) are not yet understood. For example, it seems that a moving horizons approach might be more appropriate to capture the consistent stationary sophisticated strategy that is composed of mixed actions of Section A.2.

Finally, the most interesting application of the framework presented above will be for game theory. How can players reason about the intentions and beliefs of other players, as well as their types? How can one exploit the naiveté (or sophistication) of others? What kind of equilibria are generated when (pure, or hybrid) players are pitted against each other? We hope that through this paper, we have broken the ground for such questions.

A   Comparison of naive and sophisticated decision makers in some classical decision problems

This appendix discusses a few particular decision problems to compare naive and sophisticated decision makers. In Sections A.1 and A.2, our analysis involves simple decision problems with only two states. We show that many classical examples of inconsistent behavior can already be formulated with such decision problems. In Section A.1, we tackle a class of decision problems where the decision maker faces the choice between executing an action eventually or never. We find that in the procrastination, but not in the impulsiveness problem, sophisticates might outperform naifs, depending on the exact parameters. In Section A.2 we focus on problems where an action can be chosen an arbitrary number of times, and it has both short- and long-term effects. Here, we find similar results: sophisticates can be better off than naifs in the underinvestment, but not in the binge problem. Although a lot of ground can be covered with just two states, in each of these cases sophisticates are at least as well off as naifs. To show that this is not the case for all decision problems, Section A.3 presents a decision problem with five states where naiveté dominates sophistication from a welfare perspective.

Common to the analyses is the choice of quasi-hyperbolic discounting for the utility function:

$$U^h \left(u_{t(h) \to} \right) = u_{t(h)} + \beta \sum_{t=t(h)+1}^{\infty} \delta^{t-t(h)} u_t.$$

Here, we assume $0 < \beta, \delta < 1$. The main advantages of this particular discount function are that it captures dynamic inconsistency without much technical ado, and also that the parameter β can be interpreted straightforwardly as the present-biasedness factor.

When the discount function is quasi-hyperbolic, discounting is effectively exponential from the next period onwards. Therefore, the analysis of optimal strategies boils down to two questions: under what conditions will choosing the action be optimal in the future, and under what conditions will it be optimal to perform it right away? Whenever there is a

mismatch between these two optimality conditions, the agent's incentives will change from this period to the next, and dynamic inconsistency will arise.

As our examples deal with purely naive or sophisticated decision makers, and the optimal strategies for these types are all defined to be coherent, we know that for each agent, their beliefs and intentions match whenever they choose optimal strategies. Therefore, we specify only one action for each future history. Whenever we present utility calculations, it should be clear from the agent type whether we are talking about utility based on intentions or beliefs; therefore, we drop the subscript of the utility function. Moreover, notice that our definition of a naively optimal strategy does not require perfection for the strategy – it can specify any action for future agents that will not be reached. To simplify the analysis, we therefore require our naive agents to adopt such strategies.

A.1   A single task – procrastination and impulsiveness

The following decision problem models a situation where the decision maker can perform a single task once. The state space contains only two elements: in the – initial – state ρ the task has not been chosen (yet), while in state σ it has already been performed (see Figure 9).
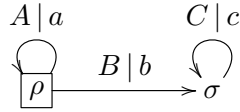
Figure 9: Single task problems.

We work with the assumption that whenever the agent is indifferent between choosing A and B, he will execute the task, choosing B. Checking first whether the agent has the incentives to execute the task in the *next* period, we get that he would choose to do so when $a \leq (1 - \delta)b + \delta c$, and would not otherwise. This yields two cases:

1. $a \leq (1 - \delta)b + \delta c$.

   The agent would execute the task in the next period, but will he execute it immediately? If he does, he gains $b + \beta \frac{\delta}{1-\delta}c$; if he does not, he gains $a + \beta \delta b + \beta \frac{\delta^2}{1-\delta}c$. This again generates two cases, where after some algebra we get:

   (a) $a \leq (1 - \beta\delta)b + \beta\delta c.$[26]

      Here, executing the task is the optimal immediate choice, as well as the optimal choice for all future periods. Both naively and sophisticatedly optimal strategies prescribe choosing B at all histories where the end-state is ρ; there is no dynamic inconsistency.

---

[26]e.g. $a = 4$, $b = 0$, $c = 20$, $\beta = \delta = 0.5$.

(b) $a > (1 - \beta\delta)b + \beta\delta c$.[27]

In this case, executing the task is optimal in the next period, but not immediately. This is the situation of procrastinating a single task.

A naively optimal strategy prescribes postponing B one period by choosing A first. A naively optimal frame would involve choosing such a strategy at every history. This implies that in the induced strategy, action B is in fact never chosen, as each naive agent keeps postponing B by one period. The induced utility over the naive frame $\tilde{f}$ is thus: $\tilde{U}_r(\tilde{f}) = a + \beta\frac{\delta}{1-\delta}a$.

Although the decision problem looks simple enough, finding the sophisticatedly optimal strategy is not straightforward. Suppose that the sophisticated agent at $\bar{h}$ believes that no future agent will choose B. If he chooses to also not do B himself, he gains $U^{\bar{h}}(AAA\ldots) = a + \beta\frac{\delta}{1-\delta}a$. On the other hand, if he takes on B, he gains $U^{\bar{h}}(B\ldots) = b + \beta\frac{\delta}{1-\delta}c$. It so happens that our two conditions thus far do not determine which of these values is larger.[28] Therefore, we need two more subcases:

i. $a + \beta\frac{\delta}{1-\delta}a \geq b + \beta\frac{\delta}{1-\delta}c$.

This means that choosing B is a best response to the belief that no future agent would choose A. Condition 1.b implies that choosing A is also a best response to the belief that in the next period B will be chosen. Furthermore, we will now show that choosing A is also a best reply to the belief that any future agent chooses B. Suppose the agent at $\bar{h}$ believes that k is the earliest period when an agent would choose B. Then, we get:

$$
\begin{aligned}
U^{\bar{h}}(AA\ldots AB\ldots) &= a + \beta\frac{\delta(1-\delta^{k-1})}{1-\delta}a + \beta\delta^k b + \beta\frac{\delta^{k+1}}{1-\delta}c \\
&= a + \beta\frac{\delta}{1-\delta}a + \beta\frac{\delta^k}{1-\delta}(-a + (1-\delta)b + \delta c) \\
&\geq a + \beta\frac{\delta}{1-\delta}a \geq b + \beta\frac{\delta}{1-\delta}c = U^{\bar{h}}(B\ldots).
\end{aligned}
$$

Note that we only get equality if both conditions 1 and 1.b.ii are satisfied with equality. It can be derived that such a scenario is impossible, given conditions 1.b and $\beta < 1$. Therefore, at least one of them is a strict inequality. This implies that A is a best response to any belief about the future, and the only sophisticatedly optimal strategy for an agent is believing and intending A for every future history. Thus, in this case, there is no difference between the induced strategy of the naively and sophisticatedly optimal frame, and the

---

[27]e.g. $a = 4$, $b = 0$, $c = 10$, $\beta = \delta = 0.5$.

[28]For instance, take $a = -1$, $b = -1$, $c = 4$, $\delta = 0.5$, $\beta = 0.5$. These values satisfy conditions 1: $a \leq (1 - \delta)b + \delta c$, 1.b: $a > (1 - \beta\delta)b + \beta\delta c$, with $a + \beta\frac{\delta}{1-\delta}a > b + \beta\frac{\delta}{1-\delta}c$. Changing only the assignment of c to $c = 6$, the parameters still satisfy conditions 1 and 1.b, but now $a + \beta\frac{\delta}{1-\delta}a < b + \beta\frac{\delta}{1-\delta}c$. Thus conditions 1.b.i/1.b.ii are independent from 1 and 1.b

induced utilities are equal.

ii. $a + \beta\frac{\delta}{1-\delta}a < b + \beta\frac{\delta}{1-\delta}c.$[29]

Choosing B is the best response for an agent that believes all future agents will choose A, and thus the strategy "all-A" is not sophisticatedly optimal. In what follows, we show that if the moment at which B is taken in the future is sufficiently far, it is optimal to choose B immediately.

Take the strategies $s_k^{\bar{h}}$ and $s_{k+1}^{\bar{h}}$, where the index $k \geq 1$ indicates the first period in which action B is taken.[30] For the utilities generated by these two strategies, we get that:

$$
\begin{aligned}
\Delta_k &= U^{\bar{h}}\left(s_k^{\bar{h}}\right) - U^{\bar{h}}\left(s_{k+1}^{\bar{h}}\right) \\
&= \left(a + \beta\frac{\delta}{1-\delta}\left(1-\delta^{k-1}\right)a + \beta\delta^k b + \beta\frac{\delta^{k+1}}{1-\delta}c\right) \\
&\quad - \left(a + \frac{\beta\delta}{1-\delta}\left(1-\delta^k\right)a + \beta\delta^{k+1}b + \beta\frac{\delta^{k+2}}{1-\delta}c\right) \\
&= \beta\delta^k\left(-a + (1-\delta)b + \delta c\right) \geq 0.
\end{aligned}
$$

So we find, by condition 1, that $U^{\bar{h}}\left(s_k^{\bar{h}}\right)$ is decreasing in k. Taking the difference between the utility generated by strategy $s_1^{\bar{h}} = AB\ldots$ and an arbitrary strategy with $k > 1$, we get:

$$
U^{\bar{h}}\left(s_1^{\bar{h}}\right) - U^{\bar{h}}\left(s_k^{\bar{h}}\right) = \sum_{i=1}^{k-1}\Delta_i = \beta\delta\frac{1-\delta^{k-1}}{1-\delta}\left(-a + (1-\delta)b + \delta c\right).
$$

Hence, we obtain:

$$
\begin{aligned}
\lim_{k\to\infty} U^{\bar{h}}\left(s_k^{\bar{h}}\right) &= U^{\bar{h}}\left(s_1^{\bar{h}}\right) - \lim_{k\to\infty}\left(U^{\bar{h}}\left(s_1^{\bar{h}}\right) - U^{\bar{h}}\left(s_k^{\bar{h}}\right)\right) \\
&= a + \beta\delta b + \beta\frac{\delta^2}{1-\delta}c - \beta\frac{\delta}{1-\delta}\left(-a + (1-\delta)b + \delta c\right) \\
&= a + \beta\frac{\delta}{1-\delta}a.
\end{aligned}
$$

We know from condition 1.b that $U^{\bar{h}}\left(s_1^{\bar{h}}\right) = U^{\bar{h}}(AB\ldots) > U^{\bar{h}}(B\ldots) = U^{\bar{h}}\left(s_0^{\bar{h}}\right)$. However, we see that $\lim_{k\to\infty} U^{\bar{h}}\left(s_k^{\bar{h}}\right) = a + \beta\frac{\delta}{1-\delta}a < b + \beta\frac{\delta}{1-\delta}c = U^{\bar{h}}(B\ldots)$. Therefore, as k grows larger, playing B will at some point become the best response to believing that A will be played until period k. In particular, there exists a smallest such $\hat{k}$, where $U^{\bar{h}}[A : s_{\hat{k}-1}^{\bar{h}}] > U^{\bar{h}}[B : s_{\hat{k}-1}^{\bar{h}}]$, but $U^{\bar{h}}[A : s_{\hat{k}}^{\bar{h}}] \leq U^{\bar{h}}[B : s_{\hat{k}}^{\bar{h}}]$.

---

[29]We ignore the possibility of mixed actions here, and instead choose to discuss it only for the underinvestment problem. Moreover, we assume that when a sophisticated decision maker can attain consistence, he will do so. For both issues, see Section A.2 condition 1.b.ii below.

[30]E.g., $s_1^{\bar{h}} = AB\ldots$, $s_2^{\bar{h}} = AAB\ldots$

Thus, an optimal strategy requires an agent to choose A, if he believes B will be chosen after at most $\hat{k} - 1$ periods, but pick B otherwise. This leads to cyclical beliefs. Table 4 represents a frame with such strategies in a compact manner: each row represent a strategy for which the row label identifies the agent's temporal distance from the root agent. Similarly, each column identifies the history that is a certain temporal distance away from the root agent.

		$t(h) - t(h_0) =$						
		0	1	2	3	4	5	...
	0	$\boxed{A}$	A	B	A	A	B	...
	1	−	$\boxed{A}$	B	A	A	B	...
	2	−	−	$\boxed{B}$	A	A	B	...
$t(\bar{h}) - t(h_0) =$	3	−	−	−	$\boxed{A}$	A	B	...
	4	−	−	−	−	$\boxed{A}$	B	...
	5	−	−	−	−	−	$\boxed{B}$	...
	...	−	−	−	−	−	−	...

Table 4: A consistent sophisticatedly optimal frame for the procrastination problem, condition 1.b.ii.

It follows that a sophisticated agent will choose a non-stationary, cyclical strategy. When it will precisely pick B depends on the beliefs of the root agent, but a consistent sophisticatedly optimal frame will definitely induce action B to be chosen sometime in the first $\hat{k}$ periods (that is, up to period $\hat{k} - 1$). It is easily checked that the expected utility of the induced strategy of the consistent sophisticatedly optimal frame will be higher than that of the naively optimal one. Thus, a sophisticated decision maker will be better off in the procrastination problem.

2. $a > (1 - \delta)b + \delta c$.

This condition means that the agent has no incentives to execute the task in the next (and subsequent) periods. Thus, his utility is $b + \beta \frac{\delta}{1-\delta}c$ if he decides to perform it immediately, while just $a + \beta \frac{\delta}{1-\delta}a$ if he does not do so (thus not performing the task at all). The two sub-cases are now:

(a) $(1 - \delta + \beta\delta)a \leq (1 - \delta)b + \beta\delta c$.[31]

Here, the agent has the incentives to do it right away given that no agent would execute the task in a later period. A naively optimal strategy simply prescribes choosing B right away, generating a utility of $\tilde{U} = b + \beta \frac{\delta}{1-\delta}c$ for the naive agent.

---

[31]e.g. $a = -4$, $b = 0$, $c = -10$, $\beta = \delta = 0.5$.

The sophisticatedly optimal strategy depends on the beliefs of the agent. Because of condition 2.a, choosing B right away is optimal if the agent believes that no future agent will take it. Now, suppose the sophisticated agent at $\bar{h}$ believes that some future agent will pick B. Let k be the earliest future period in which B will be taken, according to the strategy $s^{\bar{h}}$. We get that:

$$
\begin{aligned}
U^{\bar{h}}[B:s^{\bar{h}}] \quad &= \quad b + \tfrac{\beta\delta}{1-\delta}c \geq \tfrac{1-\delta+\beta\delta}{1-\delta}a \quad = \quad a + \tfrac{\beta\delta}{1-\delta}(a - \delta^{k-1}a + \delta^{k-1}a) \\
&> \quad a + \tfrac{\beta\delta}{1-\delta}\left((1-\delta^{k-1})a + (1-\delta)\delta^{k-1}b + \delta^k c\right) \quad = \quad U^{\bar{h}}[A:s^{\bar{h}}]
\end{aligned}
$$

We conclude that taking B right away is optimal, whatever the agent's beliefs. This means that the sophisticated decision maker cannot avoid impulsivity, and that there is no difference between the utility gained by naifs and sophisticates. In either case, if the decision maker later reflects on this decision from a non-present-biased perspective, he will regret taking B.

(b) $(1 - \delta + \beta\delta)a > (1 - \delta)b + \beta\delta c.$[32]

Now, picking A and staying in ρ is always optimal both for naifs and sophisticates, and dynamic inconsistency does not arise.

This shows that for single task problems, there is no difference in the behavior of sophisticated and naive agents for the problem of impulsiveness. In the procrastination problem, it can happen that sophistication brings no benefits, but – depending on the parameters of the problem – it might also be the case that it overperforms naiveté.

The asymmetry found between problems of impulsiveness and procrastination might look counter-intuitive at first, as the structure of the Markov decision problem looks quite similar. However, our finding is consistent with Thaler (1981), who presents some behavioral evidence that implicit discount rates are much higher for gains than for losses. If we interpret taking an impulsive action as a choice of an immediate reward, whereas taking the action that is to be procrastinated as accepting an immediate loss, our result that sophisticated decision makers can overcome the problem of procrastination are in line with their findings.

A.2 A repeated task – underinvestment and binges

The following decision problem is very similar to that in the previous one in that there are two states, with only one of them requiring a decision, whether to perform a task or not. The main difference is that the task can be executed repeatedly, as the decision maker always returns to the initial state. Figure 10 presents the parametrized version of the repeated task problem.

---

[32]e.g. $a = 4$, $b = 0$, $c = 0$, $\beta = \delta = 0.5$.

$A \mid a$

$B \mid b$
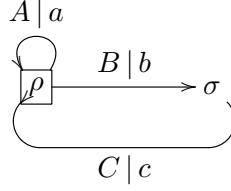
ρ → σ

$C \mid c$

Figure 10: Repeated task problem.

To simplify the analysis, we restrict our attention to strategies where the action assigned to a history h can be conditioned only on the current time and state, i.e., Markovian strategies. This allows for a very compact representation of coherent strategies, namely, as a sequence of A-s and B-s. For an agent at $\bar{h}$ the kth element in such a sequence specifies what intention-belief pair the agent assigns to period $k-1$. For instance, when we write $s^{\bar{h}} = ABAA\ldots$ it can be read as: "the agent at $\bar{h}$ (in state ρ) intends and believes A to be his current action, intends and believes action B if he is in state ρ after one period, etc." We also assume that whenever the agent is indifferent between the two actions available in state ρ, he will execute the task B.

Starting with the *next* period, the agent has the incentives to pick B whenever $a + \delta a \leq b + \delta c$. We get the following cases:

1. $a + \delta a \leq b + \delta c$.

   The agent will prefer executing B from the second period onwards. Whether he prefers to do it right away depends on the relationship between the payoff of picking A immediately, and then starting to do B, $a + \beta \frac{\delta}{1-\delta^2}(b + \delta c)$, or starting to execute B right away, $b + \beta \delta c + \beta \frac{\delta^2}{1-\delta^2}(b + \delta c)$. After some algebra, we get the following two cases:

   (a) $(1 + \delta)(a - b) \leq \beta \delta(c - b)$.[33]

      Here B is optimal whenever the state is ρ, regardless whether the agent is naive or sophisticated. Dynamic inconsistency does not arise.

   (b) $(1 + \delta)(a - b) > \beta \delta(c - b)$.[34]

      When this is the case, the decision revolves around choosing an action that brings benefits in the longer run, but these benefits are not sufficient to offset immediate costs when the agents are present-biased. Depending on the particular interpretation, this can be called a problem of slacking or underinvestment.

      Let us focus on a naive decision maker first, supposing he is at history $\bar{h}$. Our condition 1 implies that the optimal action for all futures histories is B, while condition 1.b implies that the optimal action for the current period is A. Therefore,

_____

[33]e.g. $a = 0$, $b = 0$, $c = 4$, $\beta = \delta = 0.5$.
[34]e.g. $a = 4$, $b = 0$, $c = 16$, $\beta = \delta = 0.5$.

42

the unique naively optimal strategy is $\tilde{s}^{\bar{h}} = ABBBBB\ldots$, irrespective of $\bar{h}$. The naively optimal frame is thus $\tilde{f}(h) = ABBBBB\ldots$ for all histories h. This frame is represented in a succint form in Table 5. Each row represents a set of strategies for which the row label identifies the current agent's temporal distance from the root agent. Similarly, each column identifies histories that are a certain temporal distance away from the root agent. From the representation it should be clear both that the naively optimal frame is inconsistent, and that the induced naive strategy is $\Lambda(\tilde{f}) = AAAAAA\ldots$, which would generate a utility of $U_r(\tilde{f}) = a + \beta\frac{\delta}{1-\delta}a$ for the (naive) root agent.

		$t(h) - t(h_0) =$					
		0	1	2	3	4	...
	0	$\boxed{A}$	B	B	B	B	...
	1	$-$	$\boxed{A}$	B	B	B	...
$t(\bar{h}) - t(h_0) =$	2	$-$	$-$	$\boxed{A}$	B	B	...
	3	$-$	$-$	$-$	$\boxed{A}$	B	...
	4	$-$	$-$	$-$	$-$	$\boxed{A}$	...
	...	$-$	$-$	$-$	$-$	$-$	...

Table 5: Naively optimal frame for the underinvestment problem, condition 1.b.

The analysis of the sophisticated decision maker is more complicated. Let us start with checking whether A or B is the best response of a sophisticated agent at some $\bar{h}$ whose belief is that in the next period action A will be chosen. This implies that after two periods, the current state will be ρ, so the utility generated by the rest of strategy (x) is irrelevant. The agent at $\bar{h}$ needs to compare $U^{\bar{h}}(AA\ldots) = a + \beta\delta a + x$ and $U^{\bar{h}}(BA\ldots) = b + \beta\delta c + x$. It turns out that conditions 1 and 1.b do not determine which of the two is larger.[35]

   i. $a + \beta\delta a > b + \beta\delta c$.

     This means that $U^{\bar{h}}(AA\ldots) > U^{\bar{h}}(BA\ldots)$; thus, if any A appears in the optimal strategy, it must be preceded by A-s. So, any candidate for a sophisticatedly optimal strategy has the following form: $A\ldots AB\ldots B$. However, because of condition 1.b above, $U^{\bar{h}}(ABBBBB\ldots) > U^{\bar{h}}(BBBBBB\ldots)$ and consequently there is only one candidate of such a strategy: all-A. As we have already seen, A is a best-response whenever in the next period A will be played. Therefore, for any $\bar{h}$ the sophisticatedly optimal strategy is

---

[35]To show this, take $a = -1$, $b = -1.6$, $c = 0$, $\delta = 0.75$, $\beta = 0.75$. These values satisfy conditions 1: $a + \delta a \leq b + \delta c$, and also 1.b: $(1+\delta)(a-b) > \beta\delta(c-b)$, with $a + \beta\delta a > b + \beta\delta c$. Changing only the assigment of b to $b = -1.5$, the parameters still satisfy conditions 1. and 1.b, but now $a + \beta\delta a < b + \beta\delta c$. Thus, conditions 1.b.i/1.b.ii are independent from 1 and 1.b.

$\hat{s}^{\bar{h}} = AAAAAA\dots$ While the sophisticated decision maker avoids inconsistency, his behavior will be the same as that of naive decision maker, also earning $U_r(\hat{f}) = U_r(\tilde{f}) = a + \beta\frac{\delta}{1-\delta}a$.

ii. $a + \beta\delta a \leq b + \beta\delta c$.

Now, we get $U^{\bar{h}}(BA\dots) \leq U^{\bar{h}}(AA\dots)$, so B is the best response to choosing A in the next period. What about one period before B, i.e., is $U^{\bar{h}}(ABA\dots) > U^{\bar{h}}(BBA\dots)$? It turns out that this is indeed the case. To show this, we again need to realize that since the strategy prescribes action A in the second period from the current time, the sophisticated agent knows that the current state three periods from now will be ρ, and that therefore the utility generated by the rest of the strategy (x) is irrelevant to this comparison. We get:

$$
\begin{aligned}
U^{\bar{h}}(ABA\dots) &= a + \beta\delta b + \beta\delta^2 c + x &= & \quad a - b + b + \beta\delta b + \beta\delta^2 c + x \\
&> \beta\frac{\delta}{1+\delta}(c-b) + b + \beta\delta b + \beta\delta^2 c + x \\
&= \beta\frac{\delta}{1+\delta}(c + \delta b + \delta c + \delta^2 c) + b + x = \\
&= \beta\frac{\delta}{1+\delta}(\delta b + \delta^2 c) + b + \beta\delta c + x \\
&\geq b + \beta\delta c + \beta\delta^2 a + x &= & \quad U^{\bar{h}}(BBA\dots).
\end{aligned}
$$

This implies that A is the best response to playing B next period. But as we are in condition 1.b.ii, B itself is a best response to playing A next period. So, for each $\bar{h}$ we have two candidates[36] for a sophisticatedly optimal strategy: $\hat{s}^{\bar{h}}_{AB} = ABABAB\dots$ and $\hat{s}^{\bar{h}}_{BA} = BABABA\dots$. Notice that none of these is a stationary strategy. It is then clear that a sophisticatedly optimal frame should be made up of strategies $\hat{s}_{AB}$ or $\hat{s}_{BA}$. Two consistent sophisticatedly optimal frames can be made up of these, one being represented in Table 6, the other having the entries A and B exchanged. The induced strategies are $\hat{s}_{AB}$ and $\hat{s}_{BA}$. But from the perspective of the root agent, these induced strategies generate different utilities: $U^{h_0}(\hat{s}_{AB}) = a + \beta\frac{\delta}{1-\delta^2}(b+\delta c) > b + \beta\delta c + \beta\frac{\delta^2}{1-\delta^2}(b+\delta c) = U^{h_0}(\hat{s}_{BA})$. However, it can be easily checked that both of these generate a higher payoff than the induced strategy of the naive agent.

Not only is there no unique consistent sophisticatedly optimal frame, there are numerous other *inconsistent* frames. The problem is that the beliefs of the agents are underdetermined by the best-response condition imposed by sophistication. In fact, we can create any induced strategy for the sophisticated decision maker by manipulating these agents' beliefs. For example, since

---

[36] Just like in the previous case, strategy $BBBBBB\dots$ is ruled out by condition 1.b

		$t(h) - t(h_0) =$					
		0	1	2	3	4	...
	0	$\boxed{A}$	B	A	B	A	...
	1	$-$	$\boxed{B}$	A	B	A	...
$t(\bar{h}) - t(h_0)=$	2	$-$	$-$	$\boxed{A}$	B	A	...
	3	$-$	$-$	$-$	$\boxed{B}$	A	...
	4	$-$	$-$	$-$	$-$	$\boxed{A}$	...
	...	$-$	$-$	$-$	$-$	$-$	...

Table 6: A consistent sophisticatedly optimal frame for the underinvestment problem, case 1.b.ii.

$U^{\bar{h}}(\hat{s}_{AB}) > U^{\bar{h}}(\hat{s}_{BA})$ for all $\bar{h}$, maybe each sophisticated agent believes that the agent in the next period will pick B, so that they should choose A at $\bar{h}$! We see that sophistication is no guarantee for either existence of pure stationary strategies, uniqueness of an induced sophisticated strategy, or unique payoff for the root agent. It seems that some kind of consistency of beliefs across agents is necessary to acquire a unique sophisticatedly optimal frame, and thereby a unique induced strategy. But even this might not guarantee stationarity.

One way to get a consistent sophisticatedly optimal frame of stationary strategies is to allow for randomized actions.[37] In our strategy, at each history with current state ρ, action A is played with probability p, and B with probability $1-p$. Take an agent at history $\bar{h}$, with $\omega(\bar{h}) = \rho$. First, calculate the expected utility from the next period onwards in state ρ, denoted by $x = x(p)$:

$$x = p(a + \delta x) + (1 - p)(b + \delta c + \delta^2 x).$$

We also know that the agent should be indifferent between playing A and B in $\bar{h}$, since he mixes between them. Therefore, we have:

$$a + \beta \delta x = b + \beta \delta c + \beta \delta^2 x.$$

Solving this system of equations for p, we get:

$$p = \frac{(1 + \delta)(a - b) - \beta \delta(c - b)}{\delta(1 - \beta)(a - b)}.$$

It can be shown that with the given conditions, p always falls in the $[0, 1]$ interval. By using mixed actions, we can thus construct a consistent stationary sophisticatedly optimal frame, which involves mixing between A and B with

---

[37]Note that this is a short detour from our original framework, whereby we only allowed for pure actions.

the given probability p at all histories h with $\omega(h) = \rho$. Despite the technical and conceptual difficulties, our result is a positive one for sophistication: the sophisticated decision maker can earn a higher payoff than the naive one in the underinvestment problem, given that the parameters fall into the right range.

2. $a + \delta a > b + \delta c$.

In this case, the agent will prefer picking A from the second period onwards. He will prefer to execute B immediately whenever $a + \beta\delta a \leq b + \beta\delta c$.

(a) $a + \beta\delta a \leq b + \beta\delta c.$[38]

Here, the agent has incentives to choose B once immediately, but never thereafter. A shopping binge or the consumption of an addictive substance might serve as examples for such a decision problem: present-biasedness induces the enjoyable action B, for which the costs c only need to be payed later. A naive agent thus chooses B at the outset, but intends to refrain from choosing it again later. Obviously, the frame generated by such strategies will be inconsistent, and the naive decision maker will stay addicted or keep on binging, executing B whenever in state ρ. The induced utility of this naively optimal frame is $U_r(\tilde{f}) = b + \beta\delta c + \beta\frac{\delta^2}{1-\delta^2}(b + \delta c)$. For finding sophisticatedly optimal frames, first consider an agent at history h who believes the agent in the next period in state ρ would choose A. This implies that regardless of what the agent at h chooses, the current state will be ρ two periods from now. As the payoff from that period onwards is fixed, the agent at h only needs to contrast the payoffs gained in this and the next period. By choosing A, he would gain $a + \beta\delta a$; by choosing B, he would gain $b + \beta\delta c$. With our assumption on the parameters, it is optimal for the agent at h to choose B. This means that in a sophisticated strategy, an action A is always preceded by an action B.

Now, take an agent at a history $\bar{h}$ that chooses action A at that history. There are two possibilities: either his strategy does not assign action A to any future history where the current state is ρ, or it assigns A to at least one. Suppose the first possibility; then, the strategy of the agent assigns B to all future histories where the current state is ρ. He can then choose between adopting strategy $s_A^{\bar{h}} = ABBBBB\ldots$ and strategy $s_B^{\bar{h}} = BBBBBB\ldots$. His utilities for these are:

$$
\begin{aligned}
U^{\bar{h}}\left(s_B^{\bar{h}}\right) &= b + \beta\delta c + \beta\tfrac{\delta^2}{1-\delta^2}(b + \delta c) \quad \geq \quad a + \beta\delta a + \beta\tfrac{\delta^2}{1-\delta^2}(b + \delta c) \\
&= \beta\tfrac{\delta}{1+\delta}(a + \delta a - (b + \delta c)) + a + \beta\tfrac{\delta}{1-\delta^2}(b + \delta c) \\
&> a + \beta\frac{\delta}{1 - \delta^2}(b + \delta c) \quad = \quad U^{\bar{h}}\left(s_A^{\bar{h}}\right).
\end{aligned}
$$

---

[38]e.g. $a = -4$, $b = 0$, $c = -16$, $\beta = \delta = 0.5$.

Thus, if the agent beliefs that all future agents will choose B, he should choose B himself. Therefore, one sophisticatedly optimal strategy – based on the belief that all future agents will choose B – is playing and intending to play B in all periods. The second possibility is that the agent at $\bar{h}$ believes that at least one future agent will choose A. Denote the earliest such period by $k+1$. The strategy of the agent at $\bar{h}$ can then be either $s_A^{\bar{h}} = ABB\ldots BBA\ldots$ or $s_B^{\bar{h}} = BBB\ldots BBA\ldots$, where the first and last represented action are separated by k periods of choosing B. It is clear that after period $k+1$, the agent will definitely by in state ρ – therefore, the payoffs from that point on can be ignored for the purposes of choosing an action for the agent at $\bar{h}$. We show that $U^{\bar{h}}\left(s_B^{\bar{h}}\right) > U^{\bar{h}}\left(s_A^{\bar{h}}\right)$, so that choosing B is be optimal. For simplicity, we assume k is even; the proof for k odd is analogous.

$$
\begin{aligned}
U^{\bar{h}}\left(s_B^{\bar{h}}\right) &= b + \beta\delta c + \beta\delta^2 \frac{1-\delta^k}{1-\delta^2}(b+\delta c) &\geq& \quad a + \beta\delta a + \beta\delta^2 \frac{1-\delta^k}{1-\delta^2}(b+\delta c) \\
&> a + \beta\delta\frac{1-\delta^k}{1-\delta^2}(b+\delta c) + \beta\delta^{k+1}a &=& \quad U^{\bar{h}}\left(s_A^{\bar{h}}\right).
\end{aligned}
$$

The latter inequality follows from condition 2. This shows that on the belief that some future agent will choose A, a sophisticated agent should choose B, which implies that no sophisticated strategy can include action A. As we have already shown that $BBBBB\ldots$ is a sophisticatedly optimal strategy, it is clear that sophisticates cannot outperform naifs in the binging problem.

(b) $a + \beta\delta a > b + \beta\delta c$.[39]

With this parameter configuration, picking A is always optimal for both naifs and sophisticates, and there is no issue of dynamic inconsistency.

Repeated task problems with two states involving dynamic inconsistency – i.e., underinvestment and binging problems – show a lot of similarity to single task problems. The dangers of binging behavior and impulsiveness affect sophisticates just as much as they do naifs. On the other hand, depending on the exact parameters, sophisticates can avoid underinvestment, just like they can overcome procrastination. The similarity between underinvestment and procrastination problems is even deeper, as in both cases, there might be no sophisticatedly optimal frame of stationary strategies (when we allow only for pure actions).
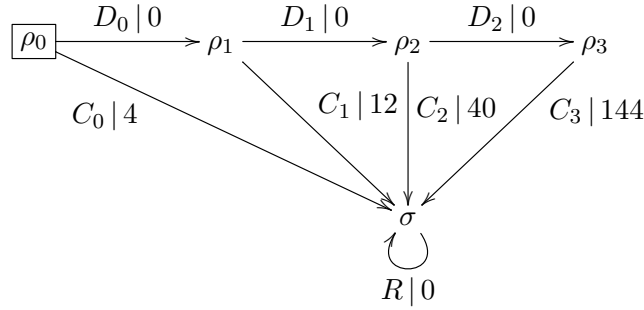
A.3 Indulgence

In the previous two decision problems, the utility induced by a sophisticatedly optimal frame was never strictly lower than that induced by a naively optimal frame. Is this always the case?

---

[39]e.g. $a = 4$, $b = 0$, $c = 0$, $\beta = \delta = 0.5$.

In this section, we review a decision problem presented by O'Donoghue and Rabin (1999a), displayed in Figure 11.

The problem is that of performing a single task in either period $0, 1, 2$ or 3 by a present-biased decision maker with $\beta = 0.5$ and $\delta = 0.5$. Think of consuming a bottle of valuable wine that gains in taste for up to three years, but then becomes undrinkable. Will one indulge in drinking it right away, or will he wait for it to fully mature? In Figure 11, "D" and "C" stand for "delay" and "consume". We ignore specifying the choice of an action for histories h with $\omega(h) = \sigma$.

$$\beta = 0.5, \delta = 0.5$$

Figure 11: A problem of indulgence.

The root agent is $h_0 = (\rho_0)$. Then, a representation of a coherent strategy for the root agent is a vector of four elements that specifies whether to open the bottle if the wine has not been consumed yet. Since only the first choice of C_i matters for induced utility, we get four relevant classes of strategies for the root agent: $S_0^{h_0} = C_0 \ldots ; S_1^{h_1} = D_0 C_1 \ldots ; S_2^{h_2} = D_0 D_1 C_2 \ldots ; S_3^{h_3} = D_0 D_1 D_2 C_3$.[40] The expected utilities calculated from the perspective of each agent for these strategy classes are presented in Table 7. For instance, $U^{h_1}(s_3^{h_1}) = 0 + \left(\frac{1}{2}\right)^2 \cdot 0 + \left(\frac{1}{2}\right)^3 \cdot 144 = 18$.

The naively optimal strategy at h_0 and h_1 is thus choosing $s_3^{h_0}$ and $s_3^{h_1}$ respectively, but at h_2 is choosing $s_2^{h_2}$. Therefore, for the naively optimal frame $\tilde{f}$, we get $\tilde{f}(h_0) \in S_3^{h_0}$, $\tilde{f}(h_1) \in S_3^{h_1}$, and $\tilde{f}(h_2) \in S_2^{h_2}$. The naive decision maker believes and intends postponing consumption right until the end, but one year before the wine fully matures, he indulges himself. The decision maker thus waits two periods, and this generates a total utility of $U_r(\tilde{f}) = 5$ for the root agent.

---

[40]In order to simplify notation, we introduce the following abbreviations: $h_0 = (\rho_0); h_1 = (\rho_0 D_0 \rho_1); h_2 = (\rho_0 D_0 \rho, D_1 \rho_2)$. Moreover, we use the following notation for strategies: strategy $s_t^{h_{t'}}$ should be read as "the agent at $h_{t'}$ intends and believes to drink the wine in period t". We will not make a strict distinction between strategies that are members of the same class – the only relevant element of a strategy is when the earliest period in which some action C_i (drinking the wine) appears. In Table 7 we are even more sloppy, and omit the reference to the agent from the indication of a strategy class.

$U^h(s)$		h		
		h_0	h_1	h_2
$s \in$	$S_0 = (C_0, \dots)$	4	–	–
	$S_1 = (D_0, C_1, \dots)$	3	12	–
	$S_2 = (D_0, D_1, C_2, \dots)$	5	10	40
	$S_3 = (D_0, D_1, D_2, C_3)$	9	18	36

Table 7: Expected utilities of coherent strategies from the various agents for the decision problem in Figure 11.

To calculate the sophisticatedly optimal frame, we resort to backward induction. The sophisticated agent at h_2 chooses $s_2^{h_2}$, for the same reasons as the naive agent, since there are no more decisions to make afterwards. Therefore, the sophisticated agent at h_1 can only consider two possibilities: $s_1^{h_1}$ or $s_2^{h_1}$. Since $U^{h_1}(s_1^{h_1}) = 12 > 10 = U^{h_1}(s_2^{h_1})$, he chooses to hasten the indulgence, since he believes (correctly) that he is unable to hold on to the end anyway. Following the same reasoning, the root agent can choose between $s_0^{h_0}$ and $s_1^{h_0}$, and opts for $s_0^{h_0}$, consuming immediately. For the sophisticatedly optimal frame $\hat{f}$, we get $\hat{f}(h_0) \in S_0^{h_0}$, $\hat{f}(h_1) \in S_1^{h_1}$ and $\hat{f}(h_2) \in S_2^{h_2}$. The sophisticated decision maker indulges himself earlier, precisely because the root agent is able to correctly reason about future agents. As the sophisticated frame generates a utility of $U_r(\tilde{f}) = 4$ for the root agent, we get $U_r(\tilde{f}) > U_r(\hat{f})$; thus, sophistication can leave somebody worse off than naiveté.

This shows that there are decision problems in which a sophisticated decision maker is worse off than a naive one. We emphasize that this is not a new result[41], but shows that our framework can easily be put to use to make such welfare comparisons between agents of various types.

Overall, by merely using the framework of quasi-hyperbolic discounting, it is possible to construct extremely simple Markov decision problems that cover most classical examples of dynamically inconsistent behavior. With these models, we can illustrate that sophistication is indeed a double-edged sword: when compared to naiveté, it might provide benefits, but it might also generate welfare losses. And sometimes, it might not influence observed behavior at all.

Moreover, our approach shows some interesting aspects of sophistication: even for decision problems with only two states, it can lead to cyclical beliefs and non-stationary strategies for all agents. For such cases, our method of representing decision makers via tables does indeed provide useful.

---

[41]Gruber and Kőszegi (2001) also find that sophistication can exascerbate drug consumption for unaddicted people. Furthermore, in a scenario that involves punishment, sophistication can actually lead to worst possible outcome (Heidhues and Kőszegi, 2009)

References

Akerlof, G. A. (1991). Procrastination and obedience. The American Economic Review, 81(2), 1–19.

Ainslie, G. W. (1974). Impulse control in pigeons. Journal of the Experimental Analysis of Behavior, 21(3), 485–489.

Ainslie, G. W. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. Psychological Bulletin, 82(4), 463.

Ainslie, G. W., and Monterosso, J. (2003). Hyperbolic discounting as a factor in addiction: A critical analysis. Choice, Behavioral Economics and Addiction, 35–61.

Ali, S. N. (2011). Learning self-control. The Quarterly Journal of Economics 126(2), 857–893.

Andersen, S., Harrison, G., Lau, M., and Rutstroem, E. (2011). Discounting behavior: a reconsideration. Working Paper. Durham University, Durham.

Angeletos, G. M., Laibson, D. I., Repetto, A., Tobacman, J., and Weinberg, S. (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. The Journal of Economic Perspectives, 15(3), 47–68.

Ariely, D., and Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. Psychological Science, 13(3), 219–224.

Asheim, G. B. (2007). Procrastination, partial naivete, and behavioral welfare analysis. No. 2007, 02. Memorandum, Department of Economics, University of Oslo.

Azfar, O. (1999). Rationalizing hyperbolic discounting. Journal of Economic Behavior & Organization, 38(2), 245–252.

Bach, C. W., and Heilmann, C. (2011). Agent connectedness and backward induction. International Game Theory Review, 13(02), 195–208.

Becker, G. S., and Murphy, K. M. (1988). A theory of rational addiction. The Journal of Political Economy, 675–700.

Benabou, R., and Pycia, M. (2002). Dynamic inconsistency and self-control: A planner-doer interpretation. Economics Letters, 77(3), 419–424.

Benhabib, J., Bisin, A., and Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. Games and Economic Behavior, 69(2), 205–223.

Blackorby, C., Nissen, D., Primont, D., and Russell, R. R. (1973). Consistent intertemporal decision making. The Review of Economic Studies, 40(2), 239–248.

Bommier, A. (2006). Uncertain lifetime and intertemporal choice: Risk aversion as a rationale for time discounting. International Economic Review, 47(4), 1223–1246.

Cowen, T. (1991). Self-constraint versus self-liberation. Ethics, 101(2), 360–373.

DellaVigna, S., and Malmendier, U. (2004). Contract design and self-control: Theory and evidence. The Quarterly Journal of Economics, 119(2), 353–402.

DellaVigna, S., and Malmendier, U. (2006). Paying not to go to the gym. The American Economic Review, 694–719.

Deweese-Boyd, I. (2012). Self-deception. The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.), Retrieved from http://plato.stanford.edu/archives/spr2012/entries/self-deception.

Eliaz, K., and Spiegler, R. (2006). Contracting with diversely naive agents. The Review of Economic Studies, 73(3), 689–714.

Elster, J. (1979). Ulysses and the sirens: Studies in rationality and irrationality. Cambridge: Cambridge University Press.

Fischer, C. (1999). Read this paper even later: Procrastination with time-inconsistent preferences. Discussion Paper 99-20, Resources for the Future.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. Journal of Economic Literature, 40(2), 351–401.

Fudenberg, D. and Levine, D. K. (1983). Subgame-perfect equilibria of finite- and infinite-horizon games. Journal of Economic Theory 31(2), 251–268.

Fudenberg, D., and Levine, D. K. (2006). A dual-self model of impulse control. The American Economic Review 96(5), 1449–1476.

Fudenberg, D., and Levine, D. K. (2012). Timing and Self-Control. Econometrica, 80(1), 1–42.

Gilpatric, S. (2003). Present-biased preferences, self-awareness and shirking in a principal-agent setting. University of Tennessee, mimeo.

Gruber, J., and Kőszegi, B. (2001). Is addiction "rational"? Theory and evidence. The Quarterly Journal of Economics, 116(4), 1261–1303.

Halevy, Y. (2008). Strotz meets Allais: Diminishing impatience and the certainty effect. The American Economic Review, 98(3), 1145–1162.

Hammond, P. J. (1976). Changing tastes and coherent dynamic choice. The Review of Economic Studies, 43(1), 159–173.

Harris, C., and Laibson, D. I. (2001). Dynamic choices of hyperbolic consumers. Econometrica, 69(4), 935–957.

Heidhues, P., and Kőszegi, B. (2009). Futile attempts at self-control. Journal of the European Economic Association 7(2-3), 423–434.

Herings, P. J. J., and Rohde, K. I. (2006). Time-inconsistent preferences in a general equilibrium model. Economic Theory, 29(3), 591–619.

Herings, P. J. J., and Rohde, K. I. (2008). On the completeness of complete markets. Economic Theory, 37(2), 171–201.

Jehiel, P., and Lilico, A. (2010). Smoking today and stopping tomorrow: a limited foresight perspective. CESifo Economic Studies 56(2), 141–164.

Howell, A. J., Watson, D. C., Powell, R, A., and Buro, K. (2006). Academic procrastination: The pattern and correlates of behavioural postponement. Personality and Individual Differences, 40(8), 1519–1530.

Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. The Quarterly Journal of Economics, 112(2), 375–406.

Kavka, G. S. (1983). The toxin puzzle. Analysis 43(1), 33–36.

Kocherlakota, N. R. (1996). Reconsideration-proofness: A refinement for infinite horizon time inconsistency. Games and Economic Behavior, 15(1), 33–54.

Laibson, D. I. (1994). Hyperbolic discounting and consumption. Doctoral dissertation, Massachusetts Institute of Technology.

Laibson, D. I. (1997). Golden eggs and hyperbolic discounting. The Quarterly Journal of Economics, 112(2), 443–478.

Laibson, D. I., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., and Akerlof, G. A. (1998). Self-control and saving for retirement. Brookings Papers on Economic Activity, 1998(1), 91–196.

Lapied, A., and Renault, O. (2012). A subjective discounted utility model. Economics Bulletin, 32(2), 1171–1179.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. Health Psychology, 24(4S), S49–S56.

Loewenstein, G., and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. The Quarterly Journal of Economics, 107(2), 573–597.

Loewenstein, G., and Thaler, R. H. (1989). Anomalies: Intertemporal choice. The Journal of Economic Perspectives, 3(4), 181–193.

McClennen, E. F. (1990). Rationality and dynamic choice. Cambridge: Cambridge University Press.

O'Donoghue, T., and Rabin, M. (1999a). Doing it now or later. American Economic Review, 89(1), 103–124.

O'Donoghue, T., and Rabin, M. (1999b). Incentives for procrastinators. The Quarterly Journal of Economics, 114(3), 769–816.

O'Donoghue, T., and Rabin, M. (2001). Choice and procrastination. The Quarterly Journal of Economics, 16(1), 121–160.

Peleg, B., and Yaari, M. E. (1973). On the existence of a consistent course of action when tastes are changing. The Review of Economic Studies, 40(3), 391–401.

Phelps, E. S., and Pollak, R A. (1968). On second-best national saving and game-equilibrium growth. The Review of Economic Studies, 35(2), 185–199.

Read, D., Frederick, S., Orsel, B., and Rahman, J. (2005). Four score and seven years from now: The date/delay effect in temporal discounting. Management Science, 51(9), 1326–1335.

Rubinstein, A. (2003). "Economics and psychology"? The case of hyperbolic discounting. International Economic Review, 44(4), 1207–1216.

Sáez-Martí, M., and Weibull, J. W. (2002). Discounting and future selves. Working Paper No. 575, IUI, The Research Institute of Industrial Economics, Stockholm.

Samuelson, P. A. (1937). A note on measurement of utility. The Review of Economic Studies, 4(2), 155–161.

Sarafidis, Y. (2004). Inter-temporal price discrimination with time inconsistent consumers. In: Econometric Society 2004 North American Summer Meetings (No. 479). Econometric Society.

Soman, D., Ainslie, G. W., Frederick, S., Li, X., Lynch, J., Moreau, P., . . . and Zauberman, G. (2005). The psychology of intertemporal discounting: Why are distant events valued differently from proximal ones? Marketing Letters, 16(3-4), 347–360.

Steel, P. (2007). The nature of procrastination: a meta-analytic and theoretical review of quintessential self-regulatory failure. Psychological bulletin, 133(1), 65–94.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. The Review of Economic Studies, 23(3), 165–180.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. Economics Letters, 8(3), 201–207.

Thaler, R. H., and Shefrin, H. M. (1981). An economic theory of self-control. The Journal of Political Economy, 89(2), 392–406.

Vieille, N., and Weibull, J. W. (2009). Multiple solutions under quasi-exponential discounting. Economic Theory, 39(3), 513–526.

Yaari, M. E. (1978). Endogenous changes in tastes: A philosophical discussion. in: Eberlein, G. and Leinfellner, W.[eds.]: Decision theory and Social Ethics. Springer Netherlands, 59–98.