Anticipating preference reversals

Yves Le Yaouanq *

November 9, 2015.

Abstract

This paper studies the consistency between a decision-maker's preferences over menus in a first period, and stochastic choices inside menus at a later date. The comparison of commitment decisions and subsequent behavior reveals whether the individual correctly anticipates future deviations from normative preferences: a sophisticated individual chooses the right commitment options, whereas a naive decision-maker overlooks some profitable opportunities. The paper provides absolute and comparative measures of naivete and shows under which conditions pessimistic menu choices can be attributed to the anticipation of decision costs. Finally, it reports the results of an experiment based on the axiomatic framework that documents a high prevalence of (partial) naivete at the individual level.

JEL codes: C91, D81, D90.

*Le Yaouanq: Toulouse School of Economics. TSE, Université de Toulouse, Manufacture des Tabacs, 21 allée de Brienne, Fr - 31000 Toulouse. E-mail: yves.le-yaouanq@m4x.org. An earlier version of this paper was circulated under the title "Anticipating Temptation". Some of the theoretical results are to be merged with "Comparative measures of naivete", a working paper independently written by David S. Ahn, Ryota Iijima and Todd Sarver. I am deeply indebted to Roland Bénabou, Christian Gollier and Jean Tirole for their guidance and encouragement. I am extremely grateful to Faruk Gul and Wolfgang Pesendorfer for their advice and insightful comments. I also thank Kai Barron, Thomas Mariotti, Stephen Morris, John Quah, Tomasz Strzalecki, Takuro Yamashita and seminar participants at Princeton University and TSE. I thank the Department of Economics of Princeton University for its hospitality and I gratefully acknowledge financial support from the Corps des Mines and the Institut Universitaire de France. This research received approval from the TSE Research Ethics Committee.

1 Introduction

The literature on time-inconsistent preferences distinguishes between two types of individuals according to their beliefs regarding their own future behavior. Sophisticated decision-makers correctly anticipate their future choices while naive individuals underestimate their propensity to deviate from their long-term goals. In the most widely used model of time-inconsistent preferences, the quasi-hyperbolic (β, δ) setting (Laibson, 1997), decision-makers are characterized by a standard discount factor δ, by a present-bias parameter β by which all future payoffs are discounted, but also by their expectation $\hat{\beta}$ of their future β: sophisticates hold correct expectations ($\hat{\beta} = \beta$), whereas naifs underestimate their present bias ($\hat{\beta} > \beta$). This literature usually shows that present bias is not an issue *per se* as soon as it is correctly forecast: a sophisticated individual can compensate for future deviations by making appropriate decisions, for instance signing optimal contracts or elaborating consistent long-term strategies. Detection and measurement of naivete about future behavior is therefore an important area of research to understand the welfare implications of time-inconsistencies in a variety of economics settings.[1]

Most existing measurements of self-control and naivete have estimated structural parameters of a (β, δ)-model and relied on aggregate tests to detect inaccurate anticipations. This paper, in contrast, proposes an axiomatic framework in which the elicitation of naivete is not tied to a specific functional form. The results are used to design an experimental method that offers several advantages with respect to existing measurements: it allows measurement of naivete at the individual level, and therefore provides information about the distribution of naivete in the population; it accounts for the decision-makers' uncertainty about their future behavior; it allows detection of partial naivete, in addition to complete sophistication and complete naivete; it is nonparametric, and encompasses all existing models of time-inconsistency. The axiomatic method makes few assumptions about the agent's preference parameters. It therefore enables understanding of the core behavioral principles linked to naivete and sophistication that would be valid in any applied model of preference reversals. It also allows an experimenter to measure the naivete of a decision-maker without specifying a fully structured model of inter-temporal preferences. As an illustration, I find in an experiment based on the theoretical framework that the majority of participants display naive anticipations, and that partial naivete is a more realistic assumption than complete

[1]See for instance Rabin and O'Donoghue (1999); Eliaz and Spiegler (2006); Heidhues and Kőszegi (2010) and Kőszegi (2014) for a survey.

naivete for most subjects.

The axiomatic framework takes as primitives the choices made by a decision-maker at two successive periods. At a first stage, called *ex ante*, the decision-maker chooses the set of options that will be available in the future, as in Kreps (1979) and Dekel et al. (2001), according to a binary preference $\succeq$ over menus. For instance, the individual chooses a menu of dishes that will be available for dinner at a future date. This choice is made when the decision-maker has certain normative preferences in mind but anticipates possible deviations, for instance because she might experience temptation at the time of actual choice. Her *ex ante* preference relation $\succeq$ is represented by a Random Strotz model (Dekel and Lipman, 2012): under this interpretation, the agent has deterministic normative preferences but is unsure about her future behavior and therefore prefers to commit to smaller menus. The relation $\succeq$ identifies the agent's beliefs about her future taste contingencies.

At a subsequent period, called *ex post*, the agent makes a stochastic choice among the available options. This decision is not necessarily consistent with the normative preferences: for instance, in the above example, the individual might be tempted to consume an appetizing but unhealthy dish instead of a healthier option. The analyst observes the random choice rule λ that determines the decision-maker's choice inside all possible menus. The *ex post* choice rule λ is represented by a Random Expected Utility model (Gul and Pesendorfer, 2006): under this interpretation, the choice in the menu is driven by the realization of an uncertain taste contingency, whose distribution is also uniquely identified from the data.

The aim of the paper is to compare the *ex ante* anticipation of taste contingencies suggested by $\succeq$ with the actual realization of *ex post* preferences identified from λ. As a first step, I provide precise definitions of sophistication and naivete in this framework. Sophisticated individuals have exactly the right model in mind about their future choices. In contrast, naive decision-makers systematically underestimate the frequency of their future deviations. These definitions generalize the conditions used in the quasi-hyperbolic model to the more general Random Strotz interpretation of preferences.

The main result characterizes the behavioral content of sophistication and naivete under the form of dynamic conditions on the pair of preferences $\{\succeq, \lambda\}$. The axioms distinguish naifs from sophisticates according to their willingness to commit: sophisticates choose exactly the right commitment options, whereas naifs mistakenly reject commitment opportunities that appear profitable in light of their *ex post* behavior. The conditions rely on the decision-maker's preference between a menu $\{p, q\}$ and a commitment device of the form $\{\kappa p + (1-\kappa)q\}$, where κ is an exogenous probability chosen by the experimenter. The singleton $\{\kappa p + (1-\kappa)q\}$ can be interpreted as a lottery that delivers p with probability κ and q with the complementary probability. Suppose that p is *ex ante* preferred to q: for instance, p is a healthy dish, whereas q is a tempting and unhealthy alternative. If κ is lower than $\lambda^{\{p,q\}}(p)$, the actual probability with which the decision-maker chooses the healthy option *ex post*, the commitment device $\{\kappa p+(1-\kappa)q\}$ is objectively inferior to the whole menu $\{p, q\}$ given the individual's *ex post* behavior. A sophisticated agent therefore rejects this commitment opportunity. This condition, labelled *No Commitment to Inferior Lotteries*, rules out pessimistic anticipations. The second axiom, *Commitment to Superior Lotteries*, rules out optimistic expectations: if $\kappa > \lambda^{\{p,q\}}(p)$, a sophisticated decision-maker accepts the commitment device. I show that extensions of these axioms to larger menus characterize sophistication. In contrast, naive agents also satisfy *No Commitment to Inferior lotteries*, but overlook some profitable opportunities and therefore violate *Commitment to Superior Lotteries*.

Comparing commitment choices with *ex post* behavior allows an experimenter to classify a subject as sophisticated or naive, but also, in the latter case, to identify the degree to which her anticipations are incorrect. Her subjective expectation regarding $\lambda^{\{p,q\}}(p)$ can be identified as the threshold κ^* that makes her indifferent between $\{p, q\}$ and $\{\kappa^* p + (1-\kappa^*)q\}$, and the distance between κ^* and $\lambda^{\{p,q\}}(p)$ provides a good measure of her degree of naivete. I generalize this observation and define a local index of naivete that measures, for each menu, the discrepancy between the decision-maker's anticipations and her actual behavior. This index is intuitively related to properties of the joint representations and can be used to define comparative measures of naivete: fixing the *ex post* behavior, a uniform increase in the index of naivete is equivalent to a downward shift in *ex ante* beliefs about the frequency of deviation, and, equivalently, to a lower demand for commitment. Fixing the preferences over menus, a uniform increase in the index of naivete is equivalent to an upward shift in *ex post* realized deviations and, equivalently, to a higher propensity to deviate from the normative goals.

The Random Strotz model is an arbitrary interpretation of commitment preferences. Indeed, there exist other representations of menu choices that are consistent with the same behavior but that suggest different beliefs over *ex post* decisions. A misinterpretation of the commitment preferences could therefore lead the analyst to reject the sophistication hypothesis by mistake. The paper studies the robustness of the method to two departures from the Strotzian interpretation. First, assuming that the decision-maker faces self-control costs *ex post*, as in the Random Gul-Pesendorfer model (Stovall, 2010), might rationalize decisions that would be classified as pessimistic in the Strotzian framework, but not choices that would be considered as naive. Second, uncertainty regarding the normative preferences reduces the willingness to commit, which is a possible confound for the identification of naivete. I describe in which situations this concern might preclude the use of the method and discuss how the analysis can be adapted to circumvent this issue.

Finally, I report the results of an experiment based on the theoretical analysis, in a setting where the normative preferences are unambiguous. While the main application of the analysis relates to temptation and self-control issues, the Random Strotz model is silent about the source of deviations from normative preferences. The experimental protocol relies on this property and studies naivete about future memory lapses.[2] Participants have the opportunity to earn a monetary prize every day within a ten day period if they remember to log in to an experimental website during the day. They earn nothing if they forget to do so. Prior to this ten sessions, their indifference threshold κ^* between this procedure and a payment rule that delivers the prize with probability κ irrespective of their behavior is elicited. Comparing κ^* with the actual probability of logon over the ten day session provides a precise measure of the difference between *ex ante* expectations and *ex post* behavior. I find that 56 percent of the participants make naive choices, while 29 percent make sophisticated decisions and only 15 percent display pessimistic expectations. The results show that naivete is not only present in the aggregate but also at the individual level for most participants, and that a majority of them is partially aware of their memory issues.

1.1 Related literature

The theoretical sections build on the literature on menu choices, started by Kreps (1979) and pursued by Dekel et al. (2001). This field provides representations for preferences over menus that allow an observer to identify the decision-maker's anticipations of her future behavior. Applications usually include preference for flexibility, reflected in a taste for larger menus, or the role of temptation, which leads the agent to prefer smaller choice sets (Gul and Pesendorfer, 2001;

[2] Memory issues are psychologically different from self-control problems, which merit specific investigation. However, testing for sophistication with respect to future behavior can be done irrespective of the force that generates deviations from normative goals.

Dekel and Lipman, 2012; Lipman and Pesendorfer, 2013). This literature is silent about the actual choice inside menus, which is usually unmodeled: the models describe how a decision-maker would behave at the *ex post* stage under the sophistication hypothesis. This paper instead augments the choice data with the actual *ex post* behavior and explicitly addresses the question of sophistication and naivete.

The first paper that provides tools to compare anticipations and realization of tastes is the work by Ahn and Sarver (2013). They study the correspondence between *ex ante* and *ex post* subjective tastes in the particular case where the agent values flexibility and not commitment.[3] The contribution of this paper is to perform the same exercise under the assumption that the decision-maker values commitment instead of flexibility, which is suited to the analysis of sophistication and naivete in the context of preference reversals. Variants of Ahn and Sarver (2013)'s main axioms are necessary but not sufficient in the Random Strotz model, as explained in section 3. Finally, Ahn et al. (2015) study a related setting in a recent paper conceived independently from this work. Some of the issues studied are similar in both papers: in particular, the analysis of sophistication and naivete in the Random Strotz model, and the definition of comparative measures of naivete. To facilitate the comparison of both papers, the differences are notified throughout the analysis. The main distinction is that this paper focuses on providing an experimental method for measuring naivete. There are also some non-overlapping parts in both papers: Ahn et al. (2015) discuss the deterministic case and some applied models, and analyze the welfare effects of commitment opportunities, whereas the present work expands the analysis to self-control preferences and contains an experimental application.

The experimental section contributes to the empirical literature that attempts to detect sophistication and naivete in the data. It also makes a methodological contribution to the literature on stochastic choice: in contrast to existing experiments that focus on group level probabilities, I elicit individual empirical choice frequencies from repeated decisions, which allows estimation of the random choice rule for all participants. The relevant work is discussed in sections 3 and 6.

The remainder of the paper is organized as follows. Section 2 introduces the primitives of the analysis and the definitions of sophistication and naivete. Section 3 provides the main representation theorem and discusses how the analysis can be used to measure naivete in experimental settings. Section 4 discusses the link between this paper and other decision-theoretic models, proposes absolute and

[3] Dean and McNeill (2015) provide an experiment based on their analysis.

comparative measures of naivete, and introduces a special class of Random Strotz models in which the main axioms take a simpler form. Section 5 expands the analysis to preferences that include self-control costs. Section 6 describes the experimental design and results. Section 7 concludes.

2 Primitives

This section describes the primitives of the model and the representations of the individual's preferences.

2.1 Objects of choice

Consider a finite set of prizes $\mathcal{Z}$, and $\Delta(\mathcal{Z})$ the set of all probability distributions on $\mathcal{Z}$, written $p, q, ...$ and called lotteries. $\Delta(\mathcal{Z})$ is endowed with the Euclidian topology, each element of $\Delta(\mathcal{Z})$ being identified with a vector of $\mathbb{R}^{|\mathcal{Z}|}$. $\mathcal{X}$ is the set of finite non-empty subsets of $\mathcal{Z}$, and elements of $\mathcal{X}$ are written $x, y, ...$ and called menus. $\mathcal{X}$ is endowed with the Hausdorff topology.

Let $\mathcal{U}$ be the set of all expected utilities on $\mathcal{Z}$. An element of $\mathcal{U}$ can be identified with a vector of $\mathbb{R}^{|\mathcal{Z}|}$. Consider the subset $\mathcal{W}$ containing all elements u of $\mathcal{U}$ that verify $\sum u(z) = 0$ and $\sum u(z)^2 = 1$. Each nonconstant expected utility can be identified with a unique element u of $\mathcal{W}$.

The behavior of a decision-maker is observed in two periods. At the *ex ante* stage, the agent has preferences over menus, as in Kreps (1979). A menu contains the options that will be available in the future. This choice is described by a preference relation $\succeq$ defined on $\mathcal{X}$, $x \succeq y$ meaning that the agent prefers to choose inside the menu x rather than in y at the later period. As usual, $\succ$ denotes the asymmetric part of $\succeq$.

At the *ex post* stage, the agent picks one element in the set. Her choice process is modeled as a random choice rule, i.e. as a function $\lambda : \mathcal{X} \to \Delta(\Delta(\mathcal{Z}))$ such that $\lambda^x(x) = 1$ for any menu $x \in \mathcal{X}$. If y is a subset of a menu x, $\lambda^x(y)$ represents the probability with which an object in y is picked when the agent chooses in x. To lighten the notation, $\lambda^x(\{p\})$ is simply written $\lambda^x(p)$.

The pair $(\succeq, \lambda)$ is the primitive of the analysis. To define sophistication and naivete in this setting, the next step consists in adding some structure into these objects in order to introduce beliefs at the *ex ante* stage and stochastic preferences at the *ex post* stage. The remainder of this section describes the representations chosen to model $\succeq$ and λ.

2.2 Random Strotz

The preference relation $\succeq$ is represented by a Random Strotz model (Dekel and Lipman, 2012).

2.2.1 Definition

Definition 2.1. The preference relation $\succeq$ admits a *Random Strotz representation* if there exists a pair (u, μ) where $u \in \mathcal{W}$ is a nontrivial expected utility and μ is a nontrivial probability distribution on $\mathcal{W}$ such that $\succeq$ is represented by

$$V(x) = \int_{\mathcal{W}} \mu(dw) \max_{p \in \mathcal{M}_w(x)} u(p) \tag{2.1}$$

where $\mathcal{M}_w(x) = \{p \in x | \forall q \in x, w(p) \geq w(q)\}$ is the set of maximizers of w in x.

To understand this representation, suppose first that the support of μ is a singleton w, and that w has a single maximizer p in the menu x. The valuation of x at the *ex ante* stage equals $u(p)$. This suggests that the agent has long-term preferences given by u at the *ex ante* stage but anticipates that she will maximize w instead of u. If w has several maximizers in x, equation 2.1 states that ties are broken in favor of u. The utility function u corresponds to the individual's normative preference, whereas the utility function w describes her actual taste at the *ex post* stage. The decision-maker displays a preference for commitment as soon as w differs from u, and does not value flexibility since the normative utility u is deterministic. Her *ex ante* preference over menus reflects the idea that she anticipates deviations from long-term goals.

The representation 2.1 adds some uncertainty about the future decision utility, keeping the normative preference certain. The utility function u is referred to as the commitment utility since it represents the preference over singleton menus. Dekel and Lipman (2012) show that the pair (u, μ) that represents $\succeq$ is unique. The measure μ identifies the decision-maker's beliefs over her future taste contingencies.

2.2.2 Decomposition

This paragraph provides a further description of the Random Strotz model. It builds on Dekel and Lipman (2012) to define two notions that play a crucial role in the analysis: the *intensity* and the *direction* of a deviation. The intensity of an *ex post* utility function describes the frequency with which deviations from long-term preferences occur. To motivate these definitions, this section first introduces a ranking defined on $\mathcal{W}$ that compares how often two *ex post* utilities deviate from the long-term preference u.

Definition 2.2. Define the order $\gg^u$ on $\mathcal{W}$ by

$$w_1 \gg^u w_2 \text{ if } u(p) > u(q), w_2(p) \geq w_2(q) \Rightarrow w_1(p) \geq w_1(q)$$

The relation $w_1 \gg^u w_2$ (to be read as "w_1 is closer to u than w_2") means that w_1 prescribes the same choice as u among pairs of lotteries at least as often as w_2. Intuitively, it means that w_2 represents a greater deviation from u than w_1. In that case, w_2 deviates from u with a *higher intensity* than w_1. As usual, $>^u$ denotes the asymmetric part of $\gg^u$.

The order $\gg^u$ is partial. The next result, due to Dekel and Lipman (2012), characterizes the sets of utility functions that can be ordered by $\gg^u$. Let $\mathcal{V} = \{v \in \mathcal{W}, u.v = 0\}$. Basic linear algebra results show that any element of $\mathcal{W} \setminus \{u, -u\}$ can be written under the form $au + \sqrt{1-a^2}v$, where $v \in \mathcal{V}$ and $a \in (-1, 1)$, and where a and v are unique.

Lemma 1. *$w_1 \gg^u w_2$ if and only if there exists $v \in \mathcal{V}$ and coefficients $a_1 \geq a_2$ such that $w_1 = a_1 u + \sqrt{1-a_1^2}v$ and $w_2 = a_2 u + \sqrt{1-a_2^2}v$.*

Fixing $v \in \mathcal{V}$, the set $\{au + \sqrt{1-a^2}v, a \in [-1, 1])\}$ can be completely ordered according to $\gg^u$, the ranking being given by the coefficients a. Conversely, two elements of $\mathcal{W}$ can be ranked if and only if they belong to such a set. If $w = au + \sqrt{1-a^2}v$, v is defined as the *direction* of the temptation w, while a measures its *intensity* (higher values of a correspond to a lower intensity).

Given these definitions, I assume two additional conditions on $\succeq$: continuity of the representation and finiteness of the set of relevant directions. Dekel and Lipman (2012) show that these properties characterize the subclass of Random Strotz models that satisfy Stovall (2010)'s axioms, in particular finiteness and continuity. These conditions are innocuous and are likely to be satisfied in all experimental settings.[4]

Given u, for any $v \in \mathcal{V}$, let $\mathcal{C}_v = \{au + \sqrt{1-a^2}v, a \in (-1, 1)\}$, and $\overline{\mathcal{C}}_v = \mathcal{C}_v \cup \{u, -u\}$ be the closure of $\mathcal{C}_v$. The set $\mathcal{W}$ can be written $\mathcal{W} = \bigcup_{v \in \mathcal{V}} \overline{\mathcal{C}}_v$. Each set $\overline{\mathcal{C}}_v$ identifies the direction of temptation v, and can be considered as a line parameterized by the intensity of temptation in that direction.

[4] The reader can refer to Dekel et al. (2001) for a discussion of these axioms.

Definition 2.3. (u, μ) is a *finite continuous-intensity Random Strotz representation* if: (i) there exists a collection of lower semi-continuous densities $\{\mu_v\}_{v\in\mathcal{V}}$ defined over $[-1, 1]$ and a probability distribution $\mu_{\mathcal{V}}$ defined on $\mathcal{V}$ such that for any measurable E,

$$\mu(E) = \int_{v\in\mathcal{V}} \mu_v(\{a \in [-1, 1] : au + \sqrt{1-a^2}v \in E\})\mu_{\mathcal{V}}(dv)$$

(ii) There exists a finite collection $\mathcal{F} \subseteq \mathcal{V}$ such that $\mu(\bigcup_{v\in\mathcal{F}} \overline{\mathcal{C}}_v) = 1$.

2.3 Random Expected Utility

This section introduces the representation of the random choice rule λ. The stochastic choice made by the agent inside menus is modeled by a Random Expected Utility representation, as axiomatized by Gul and Pesendorfer (2006). This representation allows the observer to identify a distribution of *ex post* taste contingencies, ν, with which μ can be compared. Its intuitive definition is that a particular utility function w is selected among the set $\mathcal{W}$ according to the measure ν, and the agent's choice maximizes w.

A well-known issue that arises in random utility models is that a particular w might admit multiple maximizers in the choice set. In that case, the choice prescribed by w is ambiguous. To overcome this issue, Gul and Pesendorfer (2006) and Ahn and Sarver (2013) assume that indifference is resolved according to a tie-breaking rule. The remainder of this subsection follows this procedure and describes the corresponding representation; readers who wish to skip these details might go directly to subsection 2.4.

Let $\mathbf{B}_{\mathcal{W}}$ be the Borel σ-algebra on $\mathcal{W}$ and consider $\Delta^f(\mathcal{W})$ the set of finitely additive measures over $(\mathcal{W}, \mathbf{B}_{\mathcal{W}})$. A tie-breaking rule specifies, for each $w \in \mathcal{W}$, how the choice is made in the case where w has multiple maximizers in the choice set.

Definition 2.4. A *tie-breaking rule* is a function $\tau : \mathcal{W} \to \Delta^f(\mathcal{W})$ such that, for all $x \in \mathcal{A}$ and $p \in x$:

$$\tau_w(\{v \in \mathcal{W} : \forall q \in x \setminus \{p\}, v(p) > v(q)\}) = \tau_w(\{v \in \mathcal{W} : v(p) = \max_{q\in x} v(q)\})$$

Among the set of maximizers of w in a menu x, a lottery p is chosen if the tie-breaker τ_w chooses an expected utility function $v \in \mathcal{W}$ such that p maximizes v among the maximizers of w. Hence, to be selected an element must survive a two-stage procedure: first being a maximizer of w, second being a maximizer of v among the remaining lotteries, v being chosen according to the distribution τ_w. Definition 2.4 ensures that this second comparison resolves the indifference.

Definition 2.5. λ has a *Random Expected Utility representation* if there exists a measure ν on $\mathcal{W}$ and a tie-breaking rule τ on $\mathcal{W}$ such that, for $y \subseteq x$,

$$\lambda^x(y) = \int_{w \in \mathcal{W}} \tau_w(\{v \in \mathcal{W} : \mathcal{M}_v(\mathcal{M}_w(x)) \in y\}) d\nu(w)$$

The measure ν is defined over the set of expected utility functions $\mathcal{W}$, and $\lambda^x(y)$ equals the probability with which the outcome of the two-stage process described above belongs to y. This representation suggests that a utility function w is drawn according to the measure ν, and that the agent's choice maximizes this realized preference. The measure ν identifies the realization of actual *ex post* contingencies, with which the agent's expectation, given by μ, can be compared.

Since the Random Strotz representation implicitly assumes that ties are broken in favor of u, the paper restricts attention to Random Expected Utility models that also satisfy this property: if $p \in \mathcal{M}_w(x)$ but $p \notin \mathcal{M}_u(\mathcal{M}_w(x))$, p is chosen with probability zero by the tie-breaking rule.

Assumption 1. $\forall x \in \mathcal{A}, \forall w \in \mathcal{W}, \tau_w(\{v : \mathcal{M}_v(\mathcal{M}_w(x)) \not\subset \mathcal{M}_u(\mathcal{M}_w(x))\}) = 0.$

An alternative possibility would be to include this property in the definition of sophistication. This would introduce some cumbersome notation but the axioms and the results would be unchanged. Assumption 1 is therefore maintained for the sake of simplicity.

2.4 Sophistication and naivete

Given those primitives, the aim of the paper is to compare the metacognitive process of the agent, reflected in μ, with the actual realization of tastes, reflected in ν. Defining sophistication is straightforward: a sophisticated agent has exactly the right model in mind *ex ante* when she contemplates her future choices, which is equivalent to the equality $\mu = \nu$. Defining naivete, in contrast, requires to capture the fact that the decision-maker systematically underestimates the strength of her future deviations. To do so, the definition relies on the notion of *intensity* of deviations introduced in subsection 2.2: a naive decision-maker systematically overestimates, in a first-order stochastic sense, the extent to which *ex post* choices agree with her long-term preference u.

Definition 2.6. Consider a Random Strotz representation (u, μ) and a Random Expected Utility representation (ν, τ). (i) (u, μ, ν, τ) is *sophisticated* if $\mu = \nu$; (ii) (u, μ, ν, τ) is *naive* if for any $w \in \mathcal{W}$, $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \geq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$ with strict inequality for some w.

The aim of the paper is to provide behavioral characterizations of these definitions. Naivete is a stronger requirement than $\mu \neq \nu$: it means that the decision-maker systematically underestimates the frequency of her deviations. A notion of *pessimism* can be defined as the opposite of naivete, that is, as the condition $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \leq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$ for all w (with some strict inequality). Notice that this characterization is coarse, and some individuals cannot be unambiguously classified.[5] Since naivete has received much more attention than pessimism in the applied literature, the exposition focuses on naifs and sophisticates, but the results can easily be adapted to elicit pessimism.

3 Eliciting naivete in the Random Strotz model

This section studies behavioral characterizations of the definitions of sophistication and naivete introduced earlier. The main axioms and representation result are presented in subsection 3.1. The remainder of the section discusses the application of the elicitation procedure suggested by the axiomatic framework to measure naivete in experimental work.

3.1 Representation result

Consider two lotteries p and q such that $\{p\} \succ \{q\}$, i.e. p is preferred to q from the *ex ante* point of view. How does the decision-maker value the complete menu $\{p, q\}$? Given the Random Strotz representation, her valuation depends on her subjective probability of choosing p versus q from the whole menu at the *ex post* stage. This probability equals

$$\alpha^{\{p,q\}}(p) \stackrel{\text{def}}{=} \mu(\{w : w(p) \geq w(q)\})$$

And therefore, by equation 2.1,

$$V(\{p, q\}) = \alpha^{\{p,q\}}(p)u(p) + (1 - \alpha^{\{p,q\}}(p))u(q) \tag{3.1}$$

[5]This is the case if $\mu(\{\tilde{w} : \tilde{w} \gg^u w_1\}) > \nu(\{\tilde{w} : \tilde{w} \gg^u w_1\})$ and $\mu(\{\tilde{w} : \tilde{w} \gg^u w_2\}) < \nu(\{\tilde{w} : \tilde{w} \gg^u w_2\})$ for some utility functions w_1, w_2.

Consider now the singleton menu $\{\kappa p + (1-\kappa)q\}$, where $\kappa \in [0,1]$. This singleton can be understood as a two-stage lottery that delivers p with probability κ and q with the complementary probability.[6] Since the decision-maker has no choice to make inside the menu, her corresponding valuation equals

$$V(\{\kappa p + (1-\kappa)q\}) = \kappa u(p) + (1-\kappa)u(q) \tag{3.2}$$

Suppose now that the decision-maker faces the choice between the whole menu, $\{p, q\}$, and the commitment device $\{\kappa p + (1-\kappa)q\}$. Comparing equations 3.1 and 3.2 shows that her choice depends on the relative position of $\alpha^{\{p,q\}}(p)$ and κ: she strictly prefers the commitment device if and only if $\kappa > \alpha^{\{p,q\}}(p)$, strictly prefers the whole menu if and only if $\kappa < \alpha^{\{p,q\}}(p)$ and is indifferent if $\kappa = \alpha^{\{p,q\}}(p)$. Varying the exogenous probability κ hence allows an observer to identify $\alpha^{\{p,q\}}(p)$ and to compare this anticipation with the observed probability with which the decision-maker chooses p, equal to $\lambda^{\{p,q\}}(p)$.

The first axiom exploits this idea to detect pessimistic forecasts, whereas the second axiom elicits optimistic anticipations. Both are a direct extension of the above procedure to larger menus.

Definition 3.1. The menu y is *homogeneous* if $\{p\} \sim \{q\}$ for any $p, q \in y$.

A homogeneous menu contains options which are all equivalent from a normative point of view.

Axiom 3.1 (No Commitment to Inferior Lotteries).
If y is homogeneous, $\{p\} \succ y$, $\kappa < \lambda^{y \cup \{p\}}(p) \Rightarrow \{\kappa p + (1-\kappa)q\} \prec y \cup \{p\}$ for all $q \in y$.

The lottery $\kappa p + (1-\kappa)q$ is an *inferior lottery* as soon as $\kappa < \lambda^{y \cup \{p\}}(p)$. The term *inferior* refers to the comparison between the exogenous odds and the observed *ex post* behavior of the agent. A rational decision-maker anticipates that $\{\kappa p + (1-\kappa)q\}$ yields a lower value than the expected valuation that she derives from the whole menu, since κ is lower than her own probability of choosing p. The fact that the decision-maker discards all inferior lotteries rules out pessimistic expectations. This axiom is satisfied by sophisticated agents but also by naive individuals who hold optimistic beliefs.

Axiom 3.2 complements axiom 3.1 and detects optimistic anticipations.

[6]See Dekel et al. (2001) for a discussion of this interpretation and of the underlying independence axiom.

Axiom 3.2 (Commitment to Superior Lotteries).
If x is homogeneous, $x \succ \{q\}$, $\kappa > \lambda^{x\cup\{q\}}(x) \Rightarrow \{\kappa p + (1-\kappa)q\} \succ x \cup \{q\}$ for all $p \in x$.

The lottery $\kappa p + (1-\kappa)q$ is a *superior lottery* as soon as $\kappa > \lambda^{x\cup\{q\}}(x)$. This condition allows discriminating between sophisticated and naive agents. A naive agent underestimates her propensity to self-indulge *ex post* and fails to accept superior commitments, misguided by the wrong belief that her *ex post* choice will outperform the proposed option.

The main representation result shows that axioms 3.1 and 3.2 allow detection of sophistication and naivete.[7]

Theorem 3.1. *Suppose that* (u, μ) *represents* $\succeq$ *and that* (ν, τ) *represents* λ. *Then*
(i) (u, μ, ν, τ) *is sophisticated if and only if* $(\succeq, \lambda)$ *satisfies axioms 3.1 and 3.2;*
(ii) (u, μ, ν, τ) *is naive if and only if* $(\succeq, \lambda)$ *satisfies axiom 3.1 and violates axiom 3.2.*

3.2 Experimental method

3.2.1 Procedure

Theorem 3.1 suggests an experimental procedure to detect sophistication and naivete. For instance, given two lotteries p and q satisfying $\{p\} \succeq \{p,q\} \succeq \{q\}$ with at least one strict inequality, the experimenter can: (i) elicit $\alpha^{\{p,q\}}(p)$, identified by the indifference condition $\{\alpha^{\{p,q\}}(p)p + (1-\alpha^{\{p,q\}}(p))q\} \sim \{p,q\}$, for instance by an adapted Becker-DeGroot-Marschak mechanism or a multiple choice list; (ii) compare $\alpha^{\{p,q\}}(p)$ with $\lambda^{\{p,q\}}(p)$, the observed choice frequency; (iii) classify the decision-maker as sophisticated if $\alpha^{\{p,q\}}(p) = \lambda^{\{p,q\}}(p)$, naive if $\alpha^{\{p,q\}}(p) > \lambda^{\{p,q\}}(p)$, and pessimistic otherwise.

This procedure is employed in section 6 and offers several advantages with respect to existing methods. First, the elicitation relies on choice data only (with real stakes) and not on hypothetical questionnaires. Second, it is not tied to a particular functional form such as the quasi-hyperbolic model (Strotz, 1956; Laibson, 1997), which is encompassed as a special case (see section 4.3.3). Third, the method takes into account the decision-maker's uncertainty about her future behavior. This feature allows measurement of partial naivete, in addition to perfect sophistication and complete naivete. This contrast with most existing methods

[7] Pessimists satisfy axiom 3.2 and violate axiom 3.1: they always seize profitable commitment options but also mistakenly commit to strictly inferior lotteries.

and enables a researcher to measure the discrepancy between individuals' expectations and their own behavior, as well as to gain insight into the heterogeneity of naivete in the population. For instance, Bisin and Hyndman (2015) and Mahajan and Tarozzi (2011) consider only perfectly sophisticated or completely naive decision-makers, the former being identified by their having a positive demand for commitment. The method proposed in this paper is different. Decision-makers who do not value commitment are outside of the scope of the paper. In contrast, individuals who choose commitment devices must satisfy more stringent conditions to be classified as sophisticated. As an illustration, the experimental results reported in section 6 document that a large fraction of subjects who have a positive demand for commitment nonetheless have naive anticipations.

The best evidence provided so far on naive anticipations come from data on contract choice and actual attendance at gym clubs. Della Vigna and Malmendier (2004) document that a large fraction of individuals who buy monthly or annual memberships pay a price greater than what they would pay under a pay-per-visit contract, and that this mistake has substantial economic costs. They propose a set of possible interpretations and conclude that naivete about future attendance is the most likely psychological phenomenon. While their paper makes a clear case for the existence and importance of naive anticipations, their estimation relies on the assumption that individuals would attend the gym club with the same frequency under both contracts. In a consistent planning approach, sophisticated individuals might buy the memberships precisely to commit themselves to come more often to the gym than under the pay-per-visit plan. The unobservable commitment benefit is a possible confound for the estimated value of naivete.[8]

3.2.2 Limitations

This paragraph discusses the main limitations of the method, the settings in which its use might be problematic, and some possibilities to circumvent these issues.

Non-instrumental concerns A first caveat is that the Random Strotz model assumes that the decision-maker's preferences over menus only reflect her preferences over final consumption goods. It therefore rules out other phenomena that

[8]The authors consider this explanation and calibrate a (β, δ) model to see whether sophistication is compatible with their data. They reject this hypothesis by observing that many monthly members delay the cancellation of their plan, whereas it would be in their best interest to do so early: as a consequence, correct anticipations of future attendance can be rationalized only by assuming that people underestimate their tendency to procrastinate for the cancellation.

might influence her willingness to commit. First, individuals might have intrinsic preferences over the decision process. People might value the ability to make a decision themselves on top of their consequentialist interests, as documented by Bartling et al. (2014) and Owens et al. (2014). In contrast, other authors postulate that making decisions is undesirable, because thinking is costly (Ortoleva, 2013; Ergin and Sarver, 2010) or because controlling one's impulses requires an effort (see Baumeister et al., 2007; Gul and Pesendorfer, 2001, and section 5).

Similarly, common sense suggests that self-esteem and reputation management might prevent people from choosing commitment options, since this decision reveals the existence of their self-control issues. However, one may also argue that lapses at the *ex post* stage entail a large reputation cost, which might increase the willingness to commit. For instance, Exley and Naecker (2015) report the results of a field experiment in which the demand for commitment is higher if the choice is made in public than if this decision is kept private, which suggests that individuals signal something positive about themselves by restricting their options. All in all, the influence of non-instrumental concerns on the willingness to commit is equivocal and I leave these interesting questions for future research. In general, however, an experimenter can avoid resorting to singleton menus in the elicitation procedure, as described below.

Hidden preference for flexibility An important issue is that menu choices might hide an underlying trade-off between commitment and flexibility. Suppose, for instance, that the decision-maker has uncertain normative preferences. In the subjective state s, which happens with probability θ, her normative utility is written u_s and satisfies $u_s(p) > u_s(q)$. In state s', which occurs with probability $1-\theta$, her normative preferences verify $u_{s'}(q) > u_{s'}(p)$. Suppose that she anticipates that in state s, she actually chooses p with probability α but self-indulges and consumes q with the complementary probability. For simplicity, assume that she chooses q with probability 1 in state s'. For instance, p and q can refer to gym attendance: the decision-maker prefers to go to the gym (p) rather than staying home (q) on a normal day (in state s), but might sometimes lack the willpower to do so. At the same time, she also anticipates a state s' where it is not desirable to exercise, both from the *ex ante* and the *ex post* point of view, for instance because she is sick.

The decision-maker's *ex ante* preferences reflect a trade-off between a preference for flexibility (keeping the possibility to consume q in state s') and a desire to commit (avoid consuming q in state s). If the parameters are such that

$$\theta u_s(p) + (1-\theta)u_{s'}(p) > \theta\alpha u_s(p) + \theta(1-\alpha)u_s(q) + (1-\theta)u_{s'}(q)$$

the preference for commitment dominates the preference for flexibility, and $\{p\} \succ \{p,q\} \succ \{q\}$. In addition, denoting $u = \theta u_s + (1-\theta)u_{s'}$ the expected utility over singletons,

$$V(\{p,q\}) = xu(p) + (1-x)u(q)$$

where

$$x = \theta\alpha \frac{u_s(p) - u_s(q)}{\theta(u_s(p) - u_s(q)) + (1-\theta)(u_{s'}(p) - u_{s'}(q))}$$

Therefore, at the *ex ante* stage, the decision-maker's preferences are indistinguishable from a Random Strotz model with certain normative preference u and where the probability of choosing p *ex post* equals x: the preference for flexibility relative to the state s' remains "hidden" in the representation if this concern is dominated by the desire to commit.

The impossibility to identify the trade-off between commitment and flexibility is problematic for the application of the method. Indeed, a sophisticated individual correctly anticipates that she chooses p with probability $\theta\alpha$, but since $x > \theta\alpha$ her *ex ante* preferences verify $\{p,q\} \succ \{\theta\alpha p + (1-\theta\alpha)q\}$: she (rationally) strictly prefers the whole menu to the singleton that delivers her choices with the same probabilities. The use of axioms 3.1 and 3.2 would, in that case, incorrectly classify the decision-maker as naive.

This discussion suggests that the experimental method can be used in the way described above only in settings where the normative preference is unambiguous. This property cannot be identified by choice data such as $\succeq$ and λ only. Instead, the experimenter has to rely on her/his own judgment or to other type of data, such as the preferences expressed by the subjects in a questionnaire. Section 6 provides an example in which the deviations from long-term preferences occur because of memory lapses and, therefore, are arguably always undesirable. Other potential applications include addictions if the subjects prefer to avoid consuming the tempting good in all situations.

However, the axioms can be adapted to deal with situations where normative uncertainty seems relevant. In some environments, the normative uncertainty depends on events that are observable by the experimenter (e.g., whether the individual is sick in the above example, whether she experienced a negative shock in a financial context, etc.), in which case the method can be applied conditional on each event. If it is not the case, the normative uncertainty can sometimes be expected to be realized prior to the decision uncertainty, which allows the identification of the decision-maker's beliefs over future deviations once the normative preferences are fixed.[9] For instance, a dieter might value flexibility for her next dinner because she is uncertain about her physical activity during the day. Once she learns this information, her normative preferences become clear and the only residual uncertainty concerns her future temptations. If, in contrast, the decision-maker values flexibility because she finds it acceptable to self-indulge when the tempting dishes seem particularly tasty, the normative uncertainty relies on unobservable contingencies and is realized simultaneously to the decision uncertainty, which makes the above method inapplicable.

Finally, another possibility is to expand the domain of preferences and to impose some additional assumptions to avoid the use of singleton menus in the elicitation of preferences. For instance, suppose that the consumption goods are pairs of the form (p, a), where p is a lottery in $\mathcal{Z}$ and a a monetary prize, and that the decision-maker has additive preferences: her valuation of the singleton $\{(p, a)\}$ equals $u(p) + m(a)$ where u is her normative utility over $\mathcal{Z}$ and m is her utility function over money. Consider the choice between $\{(p, a), (q, 0)\}$ and $\{(p, 0), (q, b)\}$. The first menu offers a premium a for consuming p, whereas the second menu offers a premium b for consuming q. Assume, also, that the monetary prizes a and b are small enough to be irrelevant for the choices between p and q, more precisely that (p, a) is chosen against $(q, 0)$ exactly as often as $(p, 0)$ is chosen against $(q, 0)$, and as often as $(p, 0)$ is chosen against (q, b). Under this assumption, combined with the separability of preferences, the individual's decision reflects only her preference between the lottery that delivers a with a probability equal to her subjective probability of choosing p and the lottery that delivers b with the complementary probability. This choice allows an observer to identify the decision-maker's anticipation and to compare it to her actual behavior in order to detect naivete. Notice that this procedure involves two important and non-trivial assumptions (separability of preferences and irrelevance of small monetary stakes to *ex post* choices) that can themselves be tested.

[9] Stovall (2015) applies this idea to identify uncertain normative preferences and uncertain decision utility in a setting where the individual makes menu choices at two subsequent dates.

4 Measuring naivete in the Random Strotz model

This section discusses behavioral properties related to sophistication and naivete. It first defines notions of absolute and comparative measures of naivete that build on the tests described in section 3. It then introduces a special class of Random Strotz models relevant for most applications in which axioms 3.1 and 3.2 take a simpler form. Section 5 expands the analysis to preferences including decision costs. Readers who want to skip these additional decision-theoretic results can go directly to the experimental results in section 6.

4.1 Behavioral characterizations of naivete

***Ex ante* realism vs optimism** This paragraph shows how axioms 3.1 and 3.2 relate to more immediate but less easily testable behavioral definitions of naivete and sophistication.

Definition 4.2. $\{\succeq, \lambda\}$ is: (i) *realistic* at x if $x \sim \{\sum_{p\in x} \lambda^x(p)p\}$; (ii) *optimistic* at x if $x \succeq \{\sum_{p\in x} \lambda^x(p)p\}$.

The lottery $\{\sum_{p\in x} \lambda^x(p)p\}$ can be interpreted as an equivalent of the menu x revealed by the actual choices made by the decision-maker from x. A *realistic* agent is indifferent between a menu x and its equivalent lottery: she correctly anticipates that both menus deliver the same distribution over lotteries at the consumption stage. In contrast, an *optimistic* agent prefers a menu to its equivalent since she believes that her choices from x will be better aligned with her *ex ante* preference than they actually are.

Proposition 4.1 states that realism and optimism characterize, respectively, sophistication and naivete. Together with theorem 3.1, this result shows that an experimenter can restrict attention to tests of the form given by axioms 3.1 and 3.2 to investigate the extent to which the decision-maker's expectations are optimistic.[10]

Proposition 4.1. *(u, μ, ν, τ) is sophisticated if and only if $(\succeq, \lambda)$ is realistic at any menu. (u, μ, ν, τ) is naive if and only if $(\succeq, \lambda)$ is optimistic at any menu (and strictly optimistic for at least one menu).*

[10] Ahn et al. (2015) (independently) propose realism and optimism as their behavioral characterization of sophistication in their analysis of the Random Strotz model. This paper also shows the equivalence between realism and sophistication, but focuses on axioms 3.1 and 3.2 since these conditions are easier to test experimentally.

Consequentialism Another intuitive behavioral property, consequentialism, is implied by axioms 3.1 and 3.2. Suppose that $\lambda^{x\cup\{p\}}(p) = 0$, i.e. that p is never chosen *ex post* against options in x. Since the presence of p is irrelevant for the *ex post* choices, a sophisticated individual is indifferent *ex ante* between adding p to the menu x or not.

Axiom 4.3 (Consequentialism). $\lambda^{x\cup\{p\}}(p) = 0 \Rightarrow x \cup \{p\} \sim x$.

In contrast, naive decision-makers might violate consequentialism. For instance, an individual might be willing to pay to add a superior option p to a menu (e.g., a gym membership) even though she never chooses it *ex post*, in which case axiom 4.3 is violated.[11]

Noor (2007) provides a recursive model of immediate consumption and menu choice in which the individual becomes more tempted as the time of consumption approaches. He rules out naivete and assumes perfect sophistication which, in his case, takes the form of an axiom akin to consequentialism: $\{p\} \succ \{p, q\} \Rightarrow \mathcal{C}(\{p, q\}) = p$ where $\mathcal{C}$ denotes the deterministic choice correspondence. This property identifies sophistication in the deterministic setting, but stronger conditions are needed in the stochastic case, as theorem 3.1 shows. Kopylov (2012) relaxes the consequentialism hypothesis. He assumes that decision-makers are willing to add a superior option to their choice set, even if they anticipate that they never consume it *ex post*, due to their perfectionist impulses. In other words, the joint conditions $\{p, q\} \succ \{q\}$ and $\lambda^{\{p,q\}}(p) = 0$ are compatible with sophistication in his framework. The main difference is that this paper augments the data with an explicit model of *ex post* choice and provides a dynamic axiom to elicit naivete. In addition, the choice pattern that Kopylov (2012) interprets as an intrinsic taste for keeping (but not using) the possibility of choosing a worthy element is here considered as evidence of naivete.

[11] Ahn and Sarver (2013) show that axiom 4.3 and the converse condition, that they label *Foreseen contingencies*, are necessary and sufficient for the existence a sophisticated joint representation in the case where the decision-maker values flexibility and not commitment. Their framework is the Dekel et al. (2001) monotonic representation in which the individual anticipates a stochastic taste contingency, but there is no conflict of preferences between *ex ante* and *ex post* choices. The reason why weak axioms such as consequentialism are sufficient is that the Dekel et al. (2001) representation of flexibility-loving preferences does not identify the probabilities associated with the subjective states. In contrast, the Random Strotz model identifies the subjective probabilities, and the present paper shows that more stringent conditions are needed to obtain a sophisticated representation.

4.2 Absolute and relative measures of naivete

This section introduces a cardinal index of naivete that measures the gap between expected and realized choices, and two related comparative measures of naivete. The first measures optimistic anticipations (for a given *ex post* behavior), whereas the second measures the strength of deviations (for given *ex ante* preferences). Throughout this section, only naive preferences are considered.

4.2.1 A local index of naivete

Definition 4.3. Consider a menu x, and suppose that $(\succeq, \lambda)$ is realistic or optimistic at x. Consider $\Delta(x)$ the set of lotteries defined over x, and the subset of $\Delta(x)$ defined by

$$\mathcal{N}^{\succeq,\lambda}(x) = \{\kappa \in \Delta(x) : \{\sum_{p\in x} \lambda^x(p)p\} \prec \{\sum_{p\in x} \kappa_p p\} \prec x\}$$

The *index of naivete* of $(\succeq, \lambda)$ at x is the (normalized) volume of $\mathcal{N}^{\succeq,\lambda}(x)$:

$$N^{\succeq,\lambda}(x) = \frac{\mathcal{V}(\mathcal{N}^{\succeq,\lambda}(x))}{\mathcal{V}(\Delta(x))}$$

A lottery κ belongs to $\mathcal{N}^{\succeq,\lambda}(x)$ if the agent should objectively commit to κ instead of x but naively refuses to do so. $N^{\succeq,\lambda}(x) \in [0, 1]$ measures the stochastic distance between *ex ante* and *ex post* probabilities of choice in x, and therefore the disagreement between expectations and actual behavior. For instance, if $\{p\} \succ \{q\}$, the index at $\{p, q\}$ equals $\alpha^{\{p,q\}}(p) - \lambda^{\{p,q\}}(p)$ and measures the difference between the anticipated and the actual probability of choosing the normatively superior option.

4.2.2 Comparative measures of naivete

The first measure of naivete compares pairs of agents who have the same λ but whose commitment preferences $\succeq_1$ and $\succeq_2$ are different. Agent 1 is more naive than agent 2 if agent 1 is less willing to commit than agent 2, reflecting the fact that individual 2 has a greater awareness of her tendency to being tempted.

Definition 4.4. $(\succeq_1, \lambda)$ is *more naive* than $(\succeq_2, \lambda)$ if

$$\{p\} \succ_1 x \Rightarrow \{p\} \succ_2 x$$

This definition is equivalent to Dekel and Lipman (2012)'s definition of agent 2 being more temptation-averse than agent 1. A higher aversion to temptation is naturally interpreted as a greater degree of sophistication in the present case, since naivete is precisely characterized by the unwillingness to commit. Proposition 4.2 characterizes how this notion is reflected for the Random Strotz representations (u_1, μ_1) and (u_2, μ_2) associated with $\succeq_1$ and $\succeq_2$. A first immediate observation is that $\succeq_1$ and $\succeq_2$ have the same preference over singletons, which implies that $u_1 = u_2$. Moreover, as Dekel and Lipman (2012) show, agent 1's lower demand for commitment is equivalent to the fact that agent 1 underestimates the intensity of deviations compared to agent 2.

These two equivalent properties can also be related to the absolute measure of naivete defined in 4.3. If agent 1 is more naive than agent 2, the local indexes of naivete are ranked similarly. The reverse implication also holds under the assumption $u_1 = u_2$.

Proposition 4.2. *In addition, assuming that $u_1 = u_2$, (i) and (ii) are equivalent to: (iii) for any menu x, $N^{\succeq_1,\lambda}(x) \geq N^{\succeq_2,\lambda}(x)$.*

The second measure of naivete considers pairs of agents who have the same *ex ante* preference $\succeq$, and therefore the same beliefs about their future behavior, but whose *ex post* choices might differ. Agent 1 is more naive than agent 2 if agent 1 chooses the tempting options as least as often as agent 2 does, the definition of a tempting object being given by the common preference $\succeq$.

Definition 4.5. $(\succeq, \lambda_1)$ is *more naive* than $(\succeq, \lambda_2)$ if

$$\{p\} \succ \{q\} \text{ for all } p \in x, q \in y \Rightarrow \lambda_1^{x\cup y}(x) \leq \lambda_2^{x\cup y}(y)$$

Proposition 4.3 mirrors proposition 4.2. Agent 1 is more naive if and only if she has greater tendency to deviate from the common long-term preference, which translates into a first-order stochastic dominance of the distribution of realized tastes ν_1 over ν_2 on the intensity scale, in each direction. These two properties are also equivalent to the uniform ranking of N_1 and N_2 over all menus.[12]

[12] Ahn et al. (2015) propose a unique comparative measure of naivete that encompasses both above definitions: they posit that, if two agents have the same normative preferences, one is more naive than the other if she is both more optimistic *ex ante* and more likely to deviate *ex post*. Their definition is therefore more general, whereas the measures introduce above disentangle whether an individual is more naive than another due to her more optimistic forecasts or to her lesser propensity to behave according to her long-term plans.

Proposition 4.3. *The following statements are equivalent: (i) $(\succeq, \lambda_1)$ is more naive than $(\succeq, \lambda_2)$; (ii) For any $w \in \mathcal{W}, \nu_1(\{\tilde{w} \in \mathcal{W} : \tilde{w} \gg^u w\}) \leq \nu_2(\{\tilde{w} \in \mathcal{W} : \tilde{w} \gg^u w\}$; (iii) for any menu x, $N^{\succeq,\lambda_1}(x) \geq N^{\succeq,\lambda_2}(x)$.*

4.3 One-dimensional Random Strotz models

This section introduces a particular subclass of Random Strotz models in which axioms 3.1 and 3.2 take a simpler form. The following results show that, in most settings, the observer can restrict attention to comparisons between $\{p, q\}$ and $\{\kappa p + (1-\kappa)q\}$ instead of considering larger menus, as in axioms 3.1 and 3.2 under their most general form. This property of most Random Strotz models facilitates the application of the method to experimental settings.

The models belonging to this class are labelled *one-dimensional*, and are characterized by the fact that the support of μ admits a unique direction. This refers to a situation where the agent knows the (unique) nature of her temptation but might be uncertain about its intensity. As is shown by theorem 4.2, this class admits a natural characterization and encompasses most experimental settings. In addition, weaker versions of axioms 3.1 and 3.2 are sufficient to detect sophistication and naivete in this subclass.

4.3.1 Definition

Definition 4.6. A Random Strotz representation (u, μ) is *one-dimensional* if there exists $v \in \mathcal{V}$ such that $\mu(\overline{\mathcal{C}_v}) = 1$.

Intuitively, a representation is one-dimensional if the direction of the temptation (given by v) is unique, the only *ex ante* uncertainty concerning the intensity of the deviation in that direction. Theorem 4.2, whose proof is given in appendix B, states that axiom 4.4 characterizes one-dimensional models in the Random Strotz class.

Axiom 4.4 (Ordered Temptations).
If $\{p\} \succ \{q_1\} \sim \{q_2\}$ for any $p \in x \cup y$, then $x \cup \{q_1\} \succ x \cup \{q_2\} \Rightarrow y \cup \{q_1\} \succeq y \cup \{q_2\}$

Theorem 4.2. *Suppose that (u, μ) represents $\succeq$. Then (u, μ) is one-dimensional if and only if $\succeq$ satisfies axiom 4.4.*

Axiom 4.4 means that if q_2 is more tempting than q_1 with respect to the menu x, q_2 is also (weakly) more tempting than q_1 with respect to any other menu y. Intuitively, it means that a total order can be defined on tempting options according to their *ex post* desirability. This is the case if all temptations are appealing *ex post* for the same reason, which is likely to be the case in most experimental settings.

The following example shows when axiom 4.4 might be violated. Suppose that a decision-maker has the opportunity to commit to a schedule for her next working day, splitting her time into three activities: working, exercising and leisure. She anticipates two subjective states: one in which she is lazy to work (but enjoys exercising), and one in which she is lazy to exercise (but enjoys working). Let us write (a, b) for the option that consists in working a hours and exercising during b hours.

- According to the long-term preference, $(9, 1) \sim (5, 2) \succ (5, 0) \sim (0, 1)$.
- In state 1, $(5, 2) \succ (0, 1) \succ (9, 1) \succ (5, 0)$: the agent enjoys exercising but dislikes working.
- In state 2, $(9, 1) \succ (5, 0) \succ (5, 2) \succ (0, 1)$: the agent enjoys working but dislikes exercising.

Her preferences satisfy $\{(9, 1)\} \succ \{(9, 1), (0, 1)\}$ but $\{(9, 1)\} \sim \{(9, 1), (5, 0)\}$ since $(5, 0)$ is never chosen against $(9, 1)$. Similarly, $\{(5, 2)\} \succ \{(5, 2), (5, 0)\}$ but $\{(5, 2)\} \sim \{(5, 2), (0, 1)\}$. These conditions together violate axiom 4.4. Intuitively, $(5, 0)$ is tempting with respect to $(5, 2)$ because of the laziness to exercise in state 2, while $(0, 1)$ is tempting with respect to $(9, 1)$ because of the laziness to work in state 1.

4.3.2 Sophistication and naivete in a one-dimensional model

This paragraph explains why the tests required to characterize sophistication and naivete are easier in a one-dimensional setting than in the general framework. The proof of theorem 3.1 in Appendix A shows that attention can be restricted to homogeneous menus of size K in axioms 3.1 and 3.2, where K is the number of relevant directions in the support of μ. Therefore, in a one-dimensional setting, a weaker version of axioms 3.1 and 3.2 where the homogeneous menus are singletons is sufficient to obtain the representation result. For instance, axioms 3.1 and 3.2 can be summarized by axiom 4.5, which provides a necessary and sufficient condition for sophistication:

Axiom 4.5 (Sophistication for pairs).
If $\{p\} \succ \{q\}$, $\kappa < \lambda^{\{p,q\}}(p) \Leftrightarrow \{\kappa p + (1 - \kappa)q\} \prec \{p, q\}$

In the one-dimensional setting, this property means that an experimenter can restrict attention to commitment choices between pairs of lotteries instead of considering larger menus, which facilitates the elicitation of sophistication.

4.3.3 Quasi-hyperbolic discounting

A particular example of a one-dimensional setting is given by the (β, δ) framework (Laibson, 1997). Consider the case where the prizes are consumption streams over an infinite horizon $\{0, 1, \cdots, t, \cdots\}$. A prize c is characterized by an infinite sequence $\{c_0, \cdots, c_t, \cdots\}$ of consumption levels. Suppose that the preference over singletons u can be represented by the standard discounted-utility model, and each *ex post* taste contingency belongs to the quasi-hyperbolic discounting class, as axiomatized by Olea and Strzalecki (2014):

- There exists $\delta \in [0, 1]$ and a function $w : \mathbb{R} \to \mathbb{R}$ such that

$$u(c) = \sum_{t=1}^{+\infty} \delta^{t-1} w(c_t)$$

- There exists a measure $\mu : [0, 1] \to \mathbb{R}$ such that

$$V(x) = \int_{\beta=0}^{1} \mu(\beta) \max_{c \in \mathcal{M}_{v_\beta}(x)} u(c) d\beta$$

  where $v_\beta(c) = w(c_1) + \beta \sum_{t=1}^{+\infty} \delta^{t-1} w(c_t)$.

- There exists a measure $\nu : [0, 1] \to \mathbb{R}$ and a tie-breaker τ such that

$$\forall y \subseteq x, \lambda^x(y) = \int_{\beta=0}^{1} \tau_\beta(\{\tilde{\beta} \in [0, 1] | \mathcal{M}_{v_{\tilde{\beta}}}(\mathcal{M}_{v_\beta}(x)) \in y\}) d\nu(\beta)$$

This framework is a particular case of a one-dimensional setting in which β parameterizes the intensity of temptation. The definitions and results provided above admit the following forms:

(i) Sophistication is equivalent to the identity $\mu = \nu$. A naive joint representation is such that μ strictly dominates ν at the first-order. For instance, if μ and ν are Dirac distributions respectively on $\hat{\beta}$ and β, sophistication is equivalent to $\hat{\beta} = \beta$, while naivete is equivalent to $\hat{\beta} > \beta$.

(ii) If $\succeq_1$ is represented by μ_1 and $\succeq_2$ by μ_2, $(\succeq_1, \lambda)$ is more naive than $(\succeq_2, \lambda)$ if μ_1 dominates μ_2. If λ_1 is represented by ν_1 and λ_2 by ν_2, $(\succeq, \lambda_1)$ is more naive than $(\succeq, \lambda_2)$ if ν_2 dominates ν_1.

(iii) If $\{c\} \succ \{c, c'\}$, the index of naivete at the set $\{c, c'\}$ equals $\mu([\beta^*, 1]) - \nu([\beta^*, 1])$, where β^* is defined by $v_{\beta^*}(c) = v_{\beta^*}(c')$. Thus, the index of naivete at a pair measures the distance between the cumulative distribution functions over β measured at the switching point between the two elements of the pair.

5 Naivete in the Random Gul-Pesendorfer model

The Random Strotz model provides a possible interpretation of the behavior of a decision-maker who values smaller menus. Nevertheless, other representations of the desire for commitment are conceivable: Dekel and Lipman (2012) show that a Random Strotz model of preferences over menus is indistinguishable from a Random Gul-Pesendorfer representation that includes menu-specific decision costs, based on *ex ante* decisions only. This section therefore expands the analysis of section 3 to account for the modeler's uncertainty about the right interpretation of commitment choices. Results that are robust to both specifications are useful since they do not require to take a stand on which model is the most accurate representation of behavior.

This section first introduces the Random Gul-Pesendorfer representation (Gul and Pesendorfer, 2001; Stovall, 2010) and provides definitions of sophistication and naivete in this setting. Theorem 5.1 then shows that choices that are classified as naive under the Random Strotz interpretation of preferences are also naive under any Random Gul-Pesendorfer model: axiom 3.2 is also necessary with preferences that include self-control costs. Then, proposition 5.1 proceeds to show that choices that are classified as pessimistic under the Random Strotz interpretation might be rationalized by an appropriate Random Gul-Pesendorfer model under a condition weaker than axiom 3.1: almost any pessimistic forecast can be attributed to the anticipation of (unobservable) self-control costs.

5.1 Random Gul-Pesendorfer model

Definition 5.1. The preference relation $\succeq$ admits a *Random Gul-Pesendorfer representation* (Stovall, 2010) if there exists a nontrivial expected utility u, and a nontrivial measure η on $\mathcal{U}$ such that $\succeq$ is represented by the functional

$$V(x) = \int_{w\in\mathcal{U}} [\max_{p\in x}(u(p) + w(p)) - \max_{q\in x} w(q)]\eta(dw) \tag{5.1}$$

If η is a degenerate lottery, this definition comes down to the Gul-Pesendorfer model of temptation-driven preferences (Gul and Pesendorfer, 2001). In the subjective state w, the decision-maker trades off her long-term preference u against her short-term temptation w, choosing the element that maximizes $u + w$ in x, and incurring a self-control cost equal to $\max_{q\in x} w(q)$. In contrast to the Random Strotz model, a tempting option can lower the *ex ante* valuation of a menu even if it is never chosen, provided that its presence in the menu inflicts a self-control cost to the decision-maker. Equation 5.1 adds some uncertainty by considering that temptations are drawn according to a measure η (Stovall, 2010). The functional 5.1 does not identify *ex ante* beliefs: several measures η can rationalize the same preference (see Dekel and Lipman, 2012). Notice that definition 5.1 does not impose any normalization condition on the expected utility functions.

5.2 Naivete

Defining naivete in the Random Gul-Pesendorfer model requires a more robust definition than in section 3 to deal with the non-uniqueness of the representation. A set of equivalent Random Gul-Pesendorfer representations is classified as naive with respect to some *ex post* preference if all the representations belonging to this set are naive.

Definition 5.2. Consider a Random Gul-Pesendorfer representation (u, η) and a Random Expected Utility representation (ν, τ). (u, η, ν, τ) is *naive* if for any $w \in \mathcal{U}$, $\eta(\{\tilde{w} : (u+\tilde{w}) \gg^u w\}) \geq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$. Consider a set $\mathcal{E}$ of equivalent Random Gul-Pesendorfer models. $(\mathcal{E}, \nu, \tau)$ is *naive* if for any (u, η) belonging to $\mathcal{E}$, (u, η, ν, τ) is naive.

Theorem 5.1 is adapted from Dekel and Lipman (2012) and states that the Random Strotz representation of a preference $\succeq$ is less pessimistic about the intensity of future temptations than any of its equivalent Random Gul-Pesendorfer representations.

Theorem 5.1. *Suppose that the preference $\succeq$ admits a Random Strotz representation (u, μ) and a Random Gul-Pesendorfer (u, η). For any $w \in \mathcal{U}$,*

$$\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \leq \eta(\{\tilde{w} : (u + \tilde{w}) \gg^u w\})$$

As a consequence, if (u, μ, ν, τ) is naive then the set of Random-Gul Pesendorfer models $\mathcal{E}$ associated with $\succeq$ is such that $(\mathcal{E}, \nu, \tau)$ is naive. The same conclusion holds if (u, μ, ν, τ) is sophisticated and $\succeq$ is temptation-averse.

An immediate implication is that the degree of naivete of a naive Random Strotz model is a lower bound of the degree of naivete of all corresponding Random Gul-Pesendorfer models. Intuitively, anticipating self-control costs *ex ante* reinforces the desire to commit; therefore commitment choices that appear too optimistic in light of *ex post* choices cannot be rationalized by assuming that the decison-maker was expecting decision costs. If the behavior of a decision-maker satisfies axiom 3.1 and violates axiom 3.2, theorem 5.1 shows that this finding is sufficient to conclude that all Random Gul-Pesendorfer representations of her behavior are also identified as naive. In that case, the Random Strotz interpretation of behavior is the conservative hypothesis regarding the degree of naivete attributed to the agent's behavior. In addition, Dekel and Lipman (2012) show that, except in the trivial case where temptation is not a concern, the Random Strotz model prescribes choices that are strictly more aligned with the long-term preference than any of its equivalent Random Gul-Pesendorfer models. Therefore, a pattern of choice that is sophisticated under the Random Strotz model cannot be rationalized by a Random Gul-Pesendorfer representation.

5.3 Sophistication

Suppose now that the observed choices satisfy axiom 3.2 and violate 3.1, in which case the Random Strotz representation is classified as pessimistic. This subsection studies under which conditions it is possible to rationalize the joint preferences by a Random Gul-Pesendorfer model of commitment preferences consistent with the hypothesis of sophistication. For the sake of simplicity, this section restricts attention to the simplest possible setting with only two goods and to the case where η and ν have finite supports[13] written $\{v_s\}_{s \in S}$ and $\{w'_s\}_{s' \in S'}$. Since only two subjective states are possible, ν is simply characterized by two values $\nu(\{u\})$ and $\nu(\{-u\})$ and two tie-breakers τ_u and τ_{-u}.

The definition of sophistication must be adapted to take into account the lack of normalization in equation 5.1: a sophisticated joint representation associates every *ex ante* subjective state with a unique *ex post* taste contingency that represents the same preference and occurs with the same probability.

[13] These properties are axiomatized in Stovall (2010) and Ahn and Sarver (2013) respectively.

Definition 5.3. If $(u, \{\eta_s, v_s\}_{s\in S})$ is a finite Random Gul-Pesendorfer representation and $\{\nu_{s'}, w_{s'}, \tau_{s'}\}_{s'\in S'}$ is a finite Random Expected Utility representation, the pair $((u, \{\eta_s, v_s\}), \{\nu_{s'}, w_{s'}, \tau_{s'}\})$ is *sophisticated* if there exists a bijection $\phi : S \to S'$ such that for any s, $\eta_s = \nu_{\phi(s)}$ and $u + v_s$ and $w_{\phi(s)}$ represent the same preference.

As discussed above, consequentialism is not implied by sophistication, but an asymmetric version of this axiom is necessary. Axiom 5.1 states that options that are normatively preferred and chosen with positive probability are valued at the *ex ante* stage. It rules out the extreme pessimism of a decision-maker who incorrectly believes that she never chooses the normatively superior option *ex post*.

Axiom 5.1 (Uphill Consequentialism).
If $\{p\} \succ \{q\}$, $\lambda^{\{p,q\}}(p) > 0 \Rightarrow \{p, q\} \succ \{q\}$.

Provided that $(\succeq, \lambda)$ satisfies axiom 3.2, axioms 5.1 is the only revealed preference implication of sophistication. Proposition 5.1 states that any intermediate level of *ex ante* pessimism can be attributed to the presence of self-control costs that are not observed *ex post*.

Proposition 5.1. *Suppose that $|\mathcal{Z}| = 2$, that $\succeq$ admits a finite Random Gul-Pesendorfer representation and that λ admits a finite Random Expected Utility. $(\succeq, \lambda)$ admits a sophisticated Random Strotz or Random Gul-Pesendorfer representation if and only if $(\succeq, \lambda)$ satisfies axioms 3.2 and 5.1.*

Consider a simple experiment with two goods p and q. Suppose that the experimenter elicits the indifference threshold $\alpha^{\{p,q\}}(p)$ defined by $\{p, q\} \sim \{\alpha^{\{p,q\}}(p)p + \alpha^{\{p,q\}}(q)q\}$ and the actual choice probability $\lambda^{\{p,q\}}(p)$ of a subject. As figure 1 shows, four cases arise: (i) if $\alpha^{\{p,q\}}(p) = \lambda^{\{p,q\}}(p)$, the joint behavior is rationalizable by a Random Strotz model; (ii) if $\alpha^{\{p,q\}}(p) > \lambda^{\{p,q\}}(p)$, the choices are naive under any interpretation; (iii) if $0 < \alpha^{\{p,q\}}(p) < \lambda^{\{p,q\}}(p)$, the joint preferences can be rationalized by a Random Gul-Pesendorfer interpretation; (iv) if $\alpha^{\{p,q\}}(p) = 0$ and $\lambda^{\{p,q\}}(p) > 0$, every interpretation concludes that the subject has pessimistic beliefs. In general, pessimism is therefore almost impossible to detect, and a representation that includes unobservable self-control costs can rationalize virtually any pattern of choice that would be considered as pessimistic under the consequentialist interpretation of commitment choices.

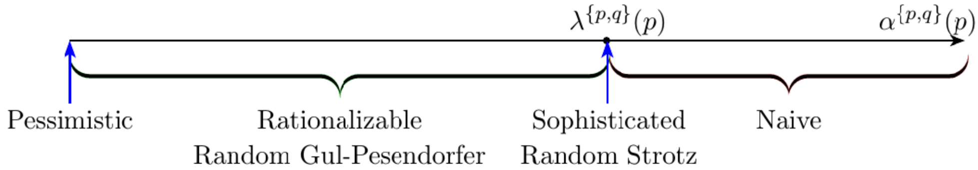

Figure 1: Classification of behavior in the two-goods case

6 Experiment

This section describes an experimental design that builds on the theoretical analysis to elicit naivete at the individual level. The experimental procedure involves a task where deviations are not due to self-control issues but to memory lapses. While both phenomena are distinct from a psychological point of view, their underlying structures are similar. Overconfident individuals either overestimate their future self-control or their future efficiency at remembering the task; in both cases, they underestimate their probability of deviating from their long-term preference, and the accuracy of their expectations can be measured by the method discussed in section 3.

6.1 Experimental procedure

6.1.1 Task

Participants have the choice between earning a monetary prize $p =$ "X dollars" or earning nothing, $q =$ "0 dollars". Deviations from long-term preferences do not arise because subjects value p and q differently at different points in time, since it is reasonable to expect them to always rank p above q. However, the *ex post* choice between p and q is made in the future and participants receive the prize p if they make an active choice at a given date, otherwise they receive the default q. A subject who forgets to claim the monetary prize p does not express preferences, but she behaves as if she ranked q above p at the time of the choice.

At the initial stage of the experiment, subjects are offered the choice between the whole choice set $\{p, q\}$ or a commitment device of the form $\{\kappa p + (1-\kappa)q\}$, where κ takes values from 0 to 1. It is reasonable to expect participants to exhibit the following preferences:

$$\{p\} \succeq \{p, q\} \succeq \{q\} \tag{6.2}$$

The commitment device $\{p\}$ always delivers the monetary prize, while the singleton $\{q\}$ delivers zero dollars with certainty. In between, the decision-maker's valuation of the choice set $\{p, q\}$ depends on her subjective probability of remembering to claim p in the future. Equation 6.2 shows that, while self-control problems and memory issues correspond to different psychological phenomena, the Random Strotz interpretation of behavior is equally suitable to both situations.

The experimental test consists in eliciting for each participant the value κ^* such that $\{p, q\}$ and $\{\kappa^* p + (1 - \kappa^*)q\}$ are valued equally at the initial stage, which is interpreted as the participant's subjective probability of remembering to claim the prize p at the *ex post* stage. Comparing κ^* with $\lambda^{\{p,q\}}(p)$, the participant's actual probability of remembering to choose the monetary reward, identifies whether the subject behaves in a sophisticated, naive or pessimistic manner.

Naivete about future prospective memory is also studied by Ericson (2011) who elicits subjects' preferences between: (i) a \$20 prize contingent on sending an e-mail during a specific slot a few months later; (ii) a sure (but lesser) prize. Assuming risk-neutrality, he shows that participants' *ex ante* choice implies that they anticipate to remember to send the e-mail with probability 76%, but the actual claim rate is only 53%. The present protocol relies on a similar task and also documents overconfidence, although the time horizon is much shorter in this one (a few days versus a few months). The contribution of this experimental section with respect to Ericson (2011)'s framework is the observation of individuals' stochastic choice at the *ex post* stage, which enables measurement of naivete at the individual level. This feature of the experimental design allows measurement of the heterogeneity in overconfidence levels. This allows the observer to determine, for instance, whether the aggregate bias is mostly due to a few subjects being extremely optimistic, or to everyone being slightly overconfident.

6.1.2 Recruitment and instructions

Subjects were recruited on Mechanical Turk, an online labor platform where individuals perform Human Intelligence Tasks on their personal computer in exchange of monetary rewards. Firms or institutions can propose tasks with a fixed payment and award bonuses depending on the quality of the worker's answers. The identity of the workers is entirely anonymous. They are identified with a personal and anonymous ID given by the website.

The lack of control over the conditions in which the subjects answer the questions might be problematic to interpret the data. For this reason, participants were asked two questions aimed at verifying their understanding of the protocol. Correctly answering both questions was necessary to receive the baseline participation fee and to be allowed to proceed with the experiment. Overall, 95 percent of the participants correctly answered both questions, suggesting that understanding issues represent a minor problem. Recent research using Mechanical Turk has shown that the quality of answers gathered on online labor markets is not significantly different from traditional laboratory experiments (Horton et al., 2011).

The task was described to the workers as an economics experiment on intertemporal decision-making. Workers who chose to participate received a link to an external website containing the experimental instructions and the answer forms. Participants were identified by means of their Mechanical Turk ID and received a personal code randomly generated to validate their participation on the Mechanical Turk platform. All payments were made on the platform by the intermediary of Amazon services. The complete experimental instructions are provided in Appendix C.

Initial session After providing informed consent, participants were informed that they would have the opportunity to earn a monetary reward every day during 10 consecutive days. Each session consisted in a 24 hours-slot during which subjects were able to earn a prize by simply signing in the experimental website with their Mechanical Turk identifier. Participants were informed that they would not receive any reminder from the experimenter and were invited to set an artificial reminder themselves. Participants were invited to write down the URL of the website. They also had the possibility to contact the experimenter at any moment through the Mechanical Turk platform to ask for the URL.

Commitment choices Subjects were offered the ooportunity to modify the payment rules for one of the ten sessions, the other nine sessions remaining unchanged. They were asked to report their preference between: (*i*) choosing later, that is, receiving the prize only conditional on visiting the website; (*ii*) being paid with probability κ, irrespective of their behavior that day. The parameter κ took 21 values for all the multiples of 5 from 0 to 100 percent. For each of these values, the participant had to choose between options (*i*) and (*ii*). One of these rows was randomly selected and the corresponding choice was implemented, thereby ensuring the incentive-compatibility of the elicitation method (Azrieli et al., 2015).

The date at which the commitment choice was relevant was selected randomly among the 10 possible session dates, and participants did not learn which date had been chosen until the end of the experiment. This procedure eliminates the effect of any private information that subjects might have regarding the evolution of their probability of remembering over time. For instance, a sophisticated participant who always remembers to log in to the website on the first day, but who always forgets to do so after that, would strictly prefer $\{p, q\}$ to $\{0.95p + 0.05q\}$ if this choice were implemented on the first day, even though her actual frequency of logon across all sessions only equals 10 percent. Implementing the payment rule at a random date makes sure that the participants should report their subjective probability of remembering to choose p averaged across the 10 sessions, which is the value that is estimated from their subsequent behavior.[14] Notice that since participants do not learn the outcome of their commitment choice before choosing the option, the Random Gul-Pesendorfer interpretation featuring decision costs is irrelevant in this situation, which allows detection of pessimistic anticipations.

Attention questions and simple checks Before submitting their choices, participants were required to answer two basic questions to verify that they read and understood the instructions. The questions were based on hypothetical scenarios: participants were asked how much they would earn depending on the row selected, their choice between the two menus, and their behavior that day. Subjects were informed that any wrong or missing answer would exclude them immediately and irrevocably from the experiment. In contrast, subjects who provided correct answers received the baseline participation fee and were allowed to proceed with the 10 regular sessions.

A simple test of understanding and rationality can also be performed by observing the commitment choices: as κ goes down, each participant should have at most one switching point from $\{\kappa p + (1-\kappa)q\}$ to $\{p, q\}$. For each subject who satisfies this criterion, the values $(\underline{\kappa}, \overline{\kappa})$ such that $\{\overline{\kappa}p + (1-\overline{\kappa})q\} \succeq \{p, q\} \succeq \{\underline{\kappa}p + (1-\underline{\kappa})q\}$ are recorded: the subjective belief of the participant is partially identified in the interval $[\underline{\kappa}, \overline{\kappa}]$.

Regular sessions The 10 regular sessions took place on the 10 days that followed the initial stage. If a participant logged in to the website during a session, the webpage displayed: a confirmation message that the visit had been registered; the history of the participant so far (dates of the sessions in which she participated,

[14] To see this, let $\Omega_t = \{0, 1\}$ be the possible events at date $t \in \{1, \cdots, 10\}$, where $\Omega_t = 1$ stands for "remembering" and $\Omega_t = 0$ for "forgetting". Consider $T \in \{1, \cdots, 10\}$ a uniform random variable over the 10 possible dates. The decision-maker's subjective probability of remembering the task at the (random) date T equals $\sum_{t=1}^{10} \mathbb{P}(T = t)\mathbb{P}(\Omega_t = 1) = \frac{1}{10}\sum_{t=1}^{10} \mathbb{P}(\Omega_t = 1)$, which is compared with the actual empirical frequency $\frac{1}{10}\sum_{t=1}^{10} \Omega_t$. Notice that this derivation does not make any assumption on the intertemporal correlations between the Ω_t.

dates of the missed sessions); the dates of the remaining sessions. For each of the 10 sessions, a dummy variable records the agent's behavior, and takes value 1 if she logged in during the session and 0 otherwise. The sum of these 10 variables λ^* yields the actual frequency with which the subject remembered to participate in the session.[15]

Payment The baseline participation fee of $1 was paid in the 12 hours that followed the initial session. The monetary prize was equal to $0.4. These values might appear very small, but they yield approximate hourly wages of respectively $20 dollars ($5 for 15 minutes of participation in total) for a subject who would pass the initial test and log in to the website every day. This amount is substantially higher than wages usually observed on the platform ($1.38 per hour as reported by Mason and Suri, 2012).

The earnings corresponding to the 10 regular sessions were paid the day after the last session under the form of a bonus on the Mechanical Turk platform. Daily payments were not provided on a regular basis because participants who had forgotten about a session would have received a payment if their commitment choice had been successful, which would have played the role of a reminder for the remaining sessions.

6.2 Results

6.2.1 Sample

A total of 189 subjects participated in the initial stage of the experiment in September 2015. 13 participants failed the attention tests and were not allowed to proceed. In addition, 5 participants provided multiple switch points and are excluded from the analysis. The final sample includes 171 participants. Average earnings amount to $3.29 per person including the participation fee.

[15]Collecting observations on stochastic choice at the individual level has been done in several experiments. In one strand of this literature, researchers ask the same question multiple times, with decoys in between (Tversky, 1969), and observe that a large fraction of subjects gives different answers. More recent work documents that participants deliberately choose randomization devices (Dwenger et al., 2014) or give different answers to the same question asked multiple times in a row (Agranov and Ortoleva, forthcoming). This suggests that the stochasticity does not necessarily arise by mistake or through variations in true preferences, but that people have an intrinsic taste for randomization. This setting rules out such phenomena. In addition, the measurement of stochastic choice at the individual level allows comparison of anticipations with actual behavior for every subject. This contrasts with Dean and McNeill (2015)'s method of estimating stochastic choice at the group level with only one choice observation per individual.

6.2.2 Raw data

This subsection provides some preliminary data comparing commitment choices with frequencies of logon.

Aggregate level The average anticipated probability of remembering to visit the website lies in the range 0.82-0.86 (std=0.24), compared with an average frequency of visit of 0.58 (std=0.43) averaged across participants and sessions. The distributions of *ex ante* and *ex post* probabilities are displayed in figure 2.

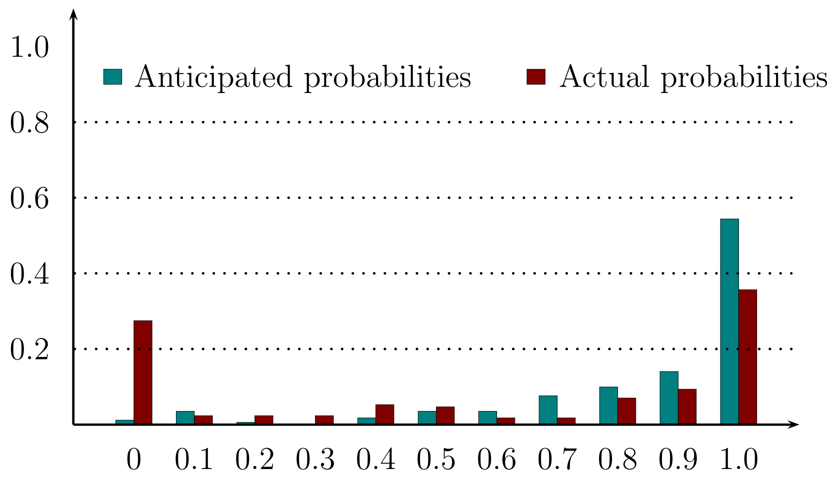

Figure 2: Distributions of anticipated and actual probabilities of visit

Individual level Each participant is characterized by a set of admissible *ex ante* beliefs $[\underline{\kappa}, \overline{\kappa}]$ and a frequency of visit λ^*. A subject's choices are classified as: naive if $\underline{\kappa} > \lambda^*$; pessimistic if $\overline{\kappa} < \lambda^*$; sophisticated otherwise. Table 1 reports the number and the fraction of participants in each category as a function of the payment. Overall, 56 percent of the subjects made naive choices, 29 percent made sophisticated decisions while 15 percent only chose inferior commitment devices. Among the subjects who made sophisticated choices, all but one (49 out of 50) exhibited extreme values of λ^* ($\lambda^* = 0$ or $\lambda^* = 1$) and correctly anticipated this. Among the subjects who missed at least one session ($\lambda^* < 1$), the prevalence of naive anticipations is substantially higher: 95 participants out of 110 (86%) had optimistic anticipations.

Measure of naivete The individual index of naivete equals $\underline{\kappa} - \lambda^*$ if the individual is naive. The average index among agents who made naive choices equals 0.51 (std=0.32), which means that these subjects on average overestimated by 51 percentage points their probability of visiting the website.

Naive choices	Sophisticated choices	Pessimistic choices	Total
95 (0.56)	50 (0.29)	26 (0.15)	171

Table 1: Number of choices per category (in parentheses, the fraction per category)

H_0 rejected (naivete)	H_0 not rejected	H_0 rejected (pessimism)	Total
60 (0.35)	110 (0.64)	1 (0.01)	171

Table 2: Test at the individual level (in parentheses, the fraction per category)

6.2.3 Statistical procedure

This section briefly discusses the estimation of participants' random choice rule from their empirical frequency.

Independent events A first possibility is to assume that participants view their behavior during the 10 sessions as independent variables drawn from the same distribution: they believe that their probability of visiting the website at any given day is independent of the day and of their behavior so far. In that case, the 10 dummy variables are independent realizations of a Bernoulli random variable, and the empirical frequency can be used as an estimate of the parameter of the Bernoulli random variable.

If the individual made a naive choice, the null hypothesis is that her probability of logon at any given date equals $\underline{\kappa}$. A one-tailed binomial test consists in computing the probability with which, under this null hypothesis of sophistication, the decision-maker's frequency of choosing p is smaller than or equal to λ^*. Similarly, for an individual who made a pessimistic choice, the test consists in computing the probability with which, under the null hypothesis of correct beliefs equal to $\overline{\kappa}$, the individual visits the website with a frequency greater than or equal to λ^*.

The results of the test are reported in table 2. Overall, the data rules out sophistication for a large fraction of agents: at the 1 percent significance level, the hypothesis is rejected for 60 optimistic agents (35 percent of the sample).[16]

[16]The number of observations (10 per individual) is relatively low. However, increasing the number of observations to gain econometric power might foster the development of habits, which makes the estimation of choice probabilities difficult (see below).

Correlated events A possible drawback of the above test is that the 10 sessions might not appear independent to the participants. For instance, a subject might believe that she will either participate in all sessions or forget entirely about the experiment, both events happening with the same probability. In that case, her *ex ante* subjective probability of visiting the website at a random date equals 0.5, but she would half of the time exhibit $\lambda^* = 0$ and be classified as naive in the above procedure. This issue is not particular to this setting: estimating individual choice probabilities requires to observe a subject making repeated decisions, and autocorrelation prevents the experimenter from considering the different choices as independent observations. The 10 observations corresponding to this individual correspond, indeed, to a unique realization of her subjective uncertainty.

If the hypothesis of independent events is relaxed, nothing can be said at the individual level except for extreme values of $(\underline{\kappa}, \overline{\kappa})$ and λ^*: for instance, choices given by $\{p\} \sim \{p, q\}$ and $\lambda^* < 1$ indicate naive anticipations, but this pattern of choice only represents 20 participants (12 percent of the sample). However, an aggregate test can be performed at the population level by observing the proportion of agents who made naive choices. For instance, the individual described above would be classified as naive with 50% chance, and as pessimist with 50% chance. If the realizations of the subjective uncertainty are independent between individuals, it is therefore possible to rule out the sophistication hypothesis by observing that a large number of subjects made a naive choice.

To obtain a conservative estimate for the proportion of naive subjects, I restrict attention to the 95 individuals who made a naive choice, and I assume that all other subjects are not naive. The 20 participants characterized by $\{p\} \sim \{p, q\}$ and $\lambda^* < 1$ are considered naive. The behavior of the remaining 75 subjects can be rationalized by an autocorrelated beliefs structure. For each of these individuals, I find the beliefs structure that rationalizes her joint behavior and that maximizes the probability with which she makes a naive choice. According to this procedure, each subject is characterized by a probability of making naive, sophisticated or pessimistic choices: for instance, the individual described above exhibits $\lambda^* = 0$ half of the time (naive choice) and $\lambda^* = 1$ half of the time (pessimistic choice). More precisely, for each individual such that $\lambda^* < \underline{\kappa}$, I assume that her *ex ante* beliefs are given by

$$\mathbb{P}(\lambda = 1) = \frac{\underline{\kappa} - \lambda^*}{1 - \lambda^*} \text{ and } \mathbb{P}(\lambda = \lambda^*) = \frac{1 - \underline{\kappa}}{1 - \lambda^*}$$

This beliefs structure rationalize the individual's behavior, since the actual empirical frequency λ^* is obtained with positive probability. In addition, this decision-maker anticipates a probability of log on equal to $\underline{\kappa}$ across the 10 sessions. It also suggests that the individual makes a naive choice with probability $\frac{1-\underline{\kappa}}{1-\lambda^*}$ and a pessimistic choice with probability $\frac{\underline{\kappa}-\lambda^*}{1-\lambda^*}$.

Under the hypothesis of sophistication, each individual is characterized by a probability of displaying a naive behavior. The number of naive choices in the population follows a Poisson-Binomial distribution whose vector of parameters is given by the individual probabilities. The proposed hypothesis is that "N individuals are naive", for a given N. To test this hypothesis, I remove from the sample the N individuals whose probability of making a naive choice is the lowest, in order to obtain a conservative estimate. The statistical tests consists in computing, given the remaining individual probabilities, the likelihood of obtaining a number of naive choices equal or greater to the number of naive choices observed in the data. I find that all hypotheses of the form "x percent of the population is naive" are rejected at the 1% significance level for any x lower than 47. These results suggest a high prevalence of naive anticipations in the population.

7 Conclusion

This paper analyzes a model of commitment decisions, augmented with *ex post* stochastic choice, to study naivete and sophistication with respect to future behavior. Its main contribution is to introduce nonparametric measures of absolute and comparative naivete that highlight the behavioral content of these phenomena while making little assumptions about the structure of preferences. An experimental design based on the axiomatic framework reveals that a large fraction of subjects are naive with respect to their future memory lapses, and that partial naivete is a better assumption than complete naivete or complete sophistication for most participants. The method can be used to understand how naivete is influenced by various parameters (experience with the task, stakes, contextual factors, etc.).

The main limitation of this approach is the possible confound created by a hidden taste for flexibility, which creates some challenges for a robust identification of naivete. Accounting for normative preference uncertainty and disentangling the flexibility-loving from the commitment-loving part of the preferences would be useful to generalize the measurement of naivete to other important settings.

Appendix A: proofs of section 3, 4 and 5

Notation

In the following, let $\mathbb{1} = (\frac{1}{|\mathcal{Z}|}, ..., \frac{1}{|\mathcal{Z}|})$ be the (scaled) unit vector. For any subset (a,b) of $[-1,1]$ and any $v \in \mathcal{V}$, let us define $\mathcal{C}_v(a,b) = \{cu + \sqrt{1-c^2}v | a < c < b\}$ and $\overline{\mathcal{C}}_v(a,b) = \{cu + \sqrt{1-c^2}v | a \leq c \leq b\}$. Let us also define $\mathcal{C}(a,b) = \bigcup_{v\in\mathcal{V}} \mathcal{C}_v(a,b)$ and $\overline{\mathcal{C}}(a,b) = \bigcup_{v\in\mathcal{V}} \overline{\mathcal{C}}_v(a,b)$.

In all this section, $\succeq$ has a finite continuous Random Strotz representation (u,μ) and λ has a Random Expected Utility representation (ν,τ). In addition, assumption 1 is satisfied.

1 Proof of theorem 3.1

1.1 Preliminary results

Lemma A.2. *$(\succeq,\lambda)$ satisfies axiom 3.1 if and only if for any $w \in \mathcal{W}$, $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \geq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$.*

Proof. Let us start with the "if" part. Take p and a homogeneous menu y such that $\{p\} \succ \{q\}$ for all $q \in y$.

Define, for $v \in \mathcal{V}$, $a(v) = \sup\{a \in [-1,1] | au(p) + \sqrt{1-a^2}v(p) \geq \max_{q\in y} au(q) + \sqrt{1-a^2}v(q)\}$ and $w(v) = a(v)u + \sqrt{1-a(v)^2}v$. If $w = au + \sqrt{1-a^2}v$, $w(p) \geq \max_{q\in y} w(q)$ is equivalent to $a \geq a(v)$, i.e. to $w \gg^u w(v)$. Thus,

$$V(y \cup \{p\}) = u(p) \int_{v\in\mathcal{V}} \mu(\{\tilde{w} : \tilde{w} \gg^u w(v)\})dv + u(q) \int_{v\in\mathcal{V}} \mu(\{\tilde{w} : w(v) \gg^u \tilde{w}\})dv$$

where q is any element of y, while

$$\lambda^{y\cup\{p\}}(p) = \int_{v\in\mathcal{V}} \nu(\{\tilde{w} : \tilde{w} \gg^u w(v)\})dv \leq \int_{v\in\mathcal{V}} \mu(\{\tilde{w} : \tilde{w} \gg^u w(v)\})dv$$

and hence $V(\{y \cup \{p\}\}) \geq \lambda^{y\cup\{p\}}(p)u(p) + \lambda^{y\cup\{p\}}(y)u(q)$

For $\kappa < \lambda^{y\cup\{p\}}(p)$, we obtain

$$\{\kappa p + (1-\kappa)q\}) \prec \{\lambda^{y\cup\{p\}}(p)p + \lambda^{y\cup\{p\}}(y)q\} \preceq y \cup \{p\}$$

which proves axiom 3.1.

The following argument proves the "only if" part by contradiction. Suppose that there exists $w \in \mathcal{W}$ such that $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) < \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$.

Write $w = au + \sqrt{1-a^2}v$, where $a \in [-1, 1]$. This condition is equivalent to $\mu(\overline{\mathcal{C}}_v(a,1)) < \nu(\overline{\mathcal{C}}_v(a,1))$. Therefore, by the continuity of μ and by theorem 10.2 of Billingsley (2012) there exists $a^* \in (a,1)$ such that $\mu(\overline{\mathcal{C}}_v(a,1)) + \mu(\overline{\mathcal{C}}(a^*,1)) < \nu(\overline{\mathcal{C}}_v(a,1))$.

Consider $\alpha > 0$, $\beta > 0$ and $\phi > 0$. Take $\tilde{w} \in \mathcal{V} \setminus \{v\}$. Since $v \neq \tilde{w}$, by the Cauchy-Schwarz inequality we have $|v.\tilde{w}| < 1$, thus it is possible to find $\gamma_{\tilde{w}}$ such that $\beta > \gamma_{\tilde{w}} v.\tilde{w}$, and $\beta v.\tilde{w} < \gamma_{\tilde{w}}$.

Define now: (i) $p = \mathbb{1} + \phi(\alpha u + \beta v)$; (ii) $q_v = \mathbb{1} + \phi\gamma_v v$ where γ_v is chosen to satisfy $au(p) + \sqrt{1-a^2}v(p) = au(q_v) + \sqrt{1-a^2}v(q_v)$; (iii) for all $\tilde{w} \in \mathcal{V}$ such that $\mu(\mathcal{C}_{\tilde{w}}) > 0$ and $\tilde{w} \neq v$, $q_{\tilde{w}} = \mathbb{1} + \phi\gamma_{\tilde{w}}\tilde{w}$. Consider finally the menu $y = \{q_v\} \cup \bigcup_{\tilde{w}\neq v}\{q_{\tilde{w}}\}$. y is finite since μ has a finite number of directions. Notice that $u(q_{\tilde{w}}) = u(q_v) = 0$ for all $\tilde{w}$, and that $u(p) > 0$.

We have $w(p) = w(q_v)$ by definition of γ_v. Moreover,

$$\begin{cases} w(p) = \phi(\alpha a + \sqrt{1-a^2}\beta) \\ w(q_{\tilde{w}}) = \phi\sqrt{1-a^2}\gamma_{\tilde{w}} v.\tilde{w} \text{ for } \tilde{w} \neq v \end{cases}$$

Since $\beta > \gamma_{\tilde{w}} v.\tilde{w}$, it is possible to pick α low enough to ensure that $w(p) \geq w(q_{\tilde{w}})$. In that case, on the set $\mathcal{C}_v$, p is chosen in the set $y \cup \{p\}$ if and only if the intensity of temptation lies in $[a,1]$[17]. Therefore $\lambda^{y\cup\{p\}}(p) \geq \nu(\overline{\mathcal{C}}_v(a,1))$.

Finally, suppose that $\hat{w} = \hat{a}u + \sqrt{1-(\hat{a})^2}\tilde{w}$, where $\tilde{w} \neq v$. We have

$$\begin{cases} \hat{w}(p) = \phi(\alpha\hat{a} + \sqrt{1-(\hat{a})^2}\beta v.\tilde{w}) \\ \hat{w}(q_{\tilde{w}}) = \phi\sqrt{1-(\hat{a})^2}\gamma_{\tilde{w}} \end{cases}$$

Since $\beta v.\tilde{w} < \gamma_{\tilde{w}}$, it is possible to pick α low enough to obtain $\hat{w}(p) < \hat{w}(q_{\tilde{w}})$ as soon as $\hat{a} < a^*$. This proves that p is chosen in $y \cup \{p\}$ at most on $\overline{\mathcal{C}}_v(a,1) \cup \overline{\mathcal{C}}(a^*,1)$.

Therefore

$$\begin{aligned} V(y \cup \{p\}) &\leq [\mu(\overline{\mathcal{C}}_v(a,1)) + \mu(\overline{\mathcal{C}}(a^*,1))]u(p) + [1 - \mu(\overline{\mathcal{C}}_v(a,1)) - \mu(\overline{\mathcal{C}}(a^*,1))]u(q) \\ &< \nu(\overline{\mathcal{C}}_v(a,1))u(p) + (1 - \nu(\overline{\mathcal{C}}_v(a,1)))u(q) \\ &\leq \lambda^{y\cup\{p\}}(p)u(p) + \lambda^{y\cup\{p\}}(y)u(q) \end{aligned}$$

It is sufficient to take $\kappa \in (\mu(\overline{\mathcal{C}}_v(a,1)) + \mu(\overline{\mathcal{C}}(a^*,1)), \nu(\overline{\mathcal{C}}_v(a,1)))$ to obtain a violation of axiom 3.1. □

[17] Remember that λ picks a maximizer of u in case of indifference, and p is the unique maximizer of u in $y \cup \{p\}$.

Lemma A.3. $(\succeq, \lambda)$ *satisfies axiom 3.2 if and only if for any* $w \in \mathcal{W}$, $\mu(\{\tilde{w} : w >^u \tilde{w}\}) \geq \nu(\{\tilde{w} : w >^u \tilde{w}\})$.

Proof. The proof is skipped. The arguments are similar to the demonstration of lemma A.2. □

1.2 Proof of necessity

Suppose that (u, μ, ν, τ) is sophisticated, i.e. that $\mu = \nu$. The conditions of lemmas A.3 and A.2 are trivially true. Therefore, axioms 3.1 and 3.2 are satisfied.

Suppose now that (u, μ, ν, τ) is naive. The condition of lemma A.2 is satisfied, therefore axiom 3.1 is valid. In addition, the continuity of $\succeq$ yields $\nu(\{u\}) = 0$.

Suppose that (u, μ, ν, τ) is strictly naive, i.e. that there exists $w \in \mathcal{W}$ such that

$$\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) > \nu(\{\tilde{w} : \tilde{w} \gg^u w\}) \tag{A.1}$$

The next step is to show that this particular w satisfies

$$\mu(\{\tilde{w} : w >^u \tilde{w}\}) < \nu(\{\tilde{w} \in \mathcal{W} | w >^u \tilde{w}\})$$

and to use lemma A.3 to conclude that axiom 3.2 is violated.

Suppose, in contrast, that

$$\mu(\{\tilde{w} \in \mathcal{W} | w >^u \tilde{w}\}) \geq \nu(\{\tilde{w} \in \mathcal{W} | w >^u \tilde{w}\}) \tag{A.2}$$

Summing A.1 and A.2 yields

$$\mu(\mathcal{C}_v) > \nu(\{-u\}) + \nu(\mathcal{C}_v) \tag{A.3}$$

Take $\tilde{w} \in \mathcal{V} \setminus \{v\}$. Since (u, μ, ν, τ) is naive, $\mu(\overline{\mathcal{C}}_{\tilde{w}}(b, 1)) \geq \nu(\overline{\mathcal{C}}_{\tilde{w}}(b, 1))$ for all $b > -1$. Taking the limit when b tends to -1 shows that

$$\mu(\mathcal{C}_{\tilde{w}}) \geq \nu(\mathcal{C}_{\tilde{w}}) \tag{A.4}$$

Integrating A.4 and summing with A.3 yields

$$\begin{aligned}\int_{\hat{w} \in \mathcal{W}} d\mu(\hat{w}) &= \mu(\mathcal{C}_v) + \int_{\tilde{w} \in \mathcal{V} \setminus \{v\}} \mu(\mathcal{C}_{\tilde{w}}) d\tilde{w} \\ &> \nu(\{-u\}) + \nu(\mathcal{C}_v) + \int_{\tilde{w} \in \mathcal{V} \setminus \{v\}} \nu(\mathcal{C}_{\tilde{w}}) d\tilde{w} = \int_{\hat{w} \in \mathcal{W}} d\nu(\hat{w})\end{aligned}$$

which is impossible since $\int_{\hat{w}\in\mathcal{W}} d\mu(\hat{w}) = \int_{\hat{w}\in\mathcal{W}} d\nu(\hat{w}) = 1$.

1.3 Proof of sufficiency

Suppose that $(\succeq, \lambda)$ satisfies axioms 3.1 and 3.2.

By lemma A.2 we have $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \geq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$ for any $w \in \mathcal{W}$, and by A.3 $\mu(\{\tilde{w} : w >^u \tilde{w}\}) \geq \mu(\{\tilde{w} : w >^u \tilde{w}\})$. We further obtain $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) = \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$ using the same argument as in the proof of necessity above. Hence, μ and ν coincide on all sets that are closed and closed by $\succeq^u$. Dekel and Lipman (2012) show that this is a sufficient condition for the measures μ and ν to coincide on all Borel sets (see their proof of theorem 1). This proves that $\mu = \nu$.

Suppose now that $(\succeq, \lambda)$ satisfies axiom 3.1 but violates axiom 3.2. We have $\mu(\{\tilde{w} : \tilde{w} \gg^u w\}) \geq \nu(\{\tilde{w} : \tilde{w} \gg^u w\})$ for every w. If this holds with equality for every w, this implies $\mu = \nu$, which contradicts the fact that $(\succeq, \lambda)$ violates axiom 3.2. Therefore there exists w such that the inequality is strict, and the representation is naive.

2 Proofs of section 4

2.1 Proof of proposition 4.1

The following paragraph proves the equivalence between realism and sophistication, the second part of the proposition being proved similarly. Suppose first that $\{\succeq, \lambda)$ is realistic. Take a homogeneous menu y, $q \in y$, p such that $\{p\} \succ \{q\}$ and $\kappa < \lambda^{y\cup\{p\}}$. We have

$$\begin{aligned} V(y \cup \{p\}) &= V(\{\lambda^{y\cup\{p\}}(p)p + \sum_{\tilde{q}\in y} \lambda^{y\cup\{p\}}(\tilde{q})\tilde{q}\}) \text{ since } (\succeq, \lambda) \text{ is realistic at } y \cup \{p\} \\ &= \lambda^{y\cup\{p\}}(p)u(p) + \lambda^{y\cup\{p\}}(y)u(q) \text{ since } y \text{ is homogenous} \\ &> \kappa u(p) + (1-\kappa)q = V(\{\kappa p + (1-\kappa)q\}) \end{aligned}$$

Hence, $\{\succeq, \lambda\}$ satisfies axiom 3.1. A similar proof shows that $\{\succeq, \lambda\}$ satisfies axiom 3.2. By theorem 3.1, $\{\succ, \lambda\}$ is sophisticated.

Suppose now that $\{\succeq, \lambda\}$ is sophisticated, and consider a finite menu x. x can be decomposed in k disjoint non-empty equivalence classes $\mathcal{E}_i, i = 1, \cdots, k$ such that $p \in \mathcal{E}_i, q \in \mathcal{E}_j$ imply $u(p) > u(q)$ if and only if $i < j$, and $u(p) = u(q)$ if $i = j$. It is therefore possible to define $\alpha^x(\mathcal{E}_i) = \mu(\{w : \mathcal{M}_u(\mathcal{M}_w(x)) \in \mathcal{E}_i\})$ the anticipated probability attached to the class $\mathcal{E}_i$[18]. Since $\mu = \nu$ and both measures break ties in favor of u, it is easy to see that $\alpha^x(\mathcal{E}_i) = \lambda^x(\mathcal{E}_i)$ for any i. Writing p_i for any element of the class $\mathcal{E}_i$, we obtain

$$V(x) = \sum_{i=1}^{k} \alpha^x(\mathcal{E}_i)u(p_i) = \sum_{i=1}^{k} \lambda^x(\mathcal{E}_i)u(p_i) = V(\{\sum_{x \in p} \lambda^x(p)p\})$$

which proves that $x \sim \{\sum_{p \in x} \lambda^x(p)p\}$. Hence, $\{\succeq, \lambda\}$ is realistic.

2.2 Proof of proposition 4.2

$(i) \Leftrightarrow (ii)$ This result is proved by Dekel and Lipman (2012) in their Theorem 4, p. 1284.

$(i) \Rightarrow (iii)$ Consider a menu x, and a lottery $\kappa \in \Delta(x)$. Since $(\succeq_1, \lambda)$ is more naive than $(\succeq_2, \lambda)$,

$$\{\sum_{p \in x} \kappa_p p\} \prec_1 x \Rightarrow \{\sum_{p \in x} \kappa_p p\} \prec_2 x$$

Therefore $\mathcal{N}^{\succeq_2, \lambda}(x) \subseteq \mathcal{N}^{\succeq_1, \lambda}(x)$, which implies $N^{\succeq_2, \lambda}(x) \leq N^{\succeq_1, \lambda}(x)$.

$(iii) \Rightarrow (i)$ For $i = 1, 2$, let us write V_i for the functional associated with $\succeq_i$ and u the (common) normative utility function. Consider a menu x. A lottery κ on $\Delta(x)$ belongs to $\mathcal{N}^{\succeq_i, \lambda}(x)$ if and only if

$$\sum_{p \in x} \lambda^x(p)u(p) < \sum_{p \in x} \kappa_p u(p) < V_i(x)$$

Hence, from $N^{\succeq_1, \lambda}(x) \geq N^{\succeq_2, \lambda}(x)$ we obtain $V_1(x) \geq V_2(x)$. Now, for any lottery p we have

$$\{p\} \succ_1 x \Leftrightarrow u(p) > V_1(x) \Rightarrow u(p) > V_2(x) \Leftrightarrow \{p\} \succ_2 x$$

which proves that $(\succeq_1, \lambda)$ is more naive than $(\succeq_2, \lambda)$.

2.3 Proof of proposition 4.3

The proof is skipped. It relies on the same arguments as proposition 4.2.

[18]The representation does not specify how the choice is made inside a class $\mathcal{E}_i$ between two options that are equally valued at the *ex post* stage, but this choice is irrelevant here.

3 Proofs of section 5

3.1 Proof of theorem 5.1

The proof is given by Dekel and Lipman (2012) (see their theorem 5 p. 1286).

3.2 Proof of proposition 5.1

In the two goods case, only two generic utilities $u \in \mathcal{W}$ exist: $u = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ and $-u$. Therefore each subjective utility can be written $v_s = \alpha_s u$ where $\alpha_s \notin \{0,1\}$. A Random Gul-Pesendorfer representation is simply written

$$V(x) = \sum_{s \in S} \eta_s [\max_{p \in x} (1+\alpha_s) u(p) - \max_{q \in x} \alpha_s u(q)] \tag{A.5}$$

The decision utility $(1+\alpha_s)u$ coincides with u if and only if $\alpha_s > -1$. A Random Expected Utility model ν simply specifies the weights ν_u and ν_{-u} attached to the states u and $-u$ respectively.

Proof of necessity If$(\succeq, \lambda)$ admits a sophisticated Random Strotz representation, necessity of axioms 3.2 and 5.1 are established in section 3. Suppose that $(\succeq, \lambda)$ admits a sophisticated Random Gul-Pesendorfer representation, i.e. that there exists a representation A.5 of $\succeq$ such that $\sum_{s \in S, \alpha_s > -1} \eta_s = \nu_u$. Equation A.5 yields

$$V(\{p,q\}) = \sum_{\alpha_s > 0} \eta_s u(p) + \sum_{-1 < \alpha_s < 0} \eta_s [u(p) + \alpha_s (u(p) - u(q))] + \sum_{\alpha_s < -1} \eta_s u(q) \tag{A.6}$$

If there exists p, q such that $\{p\} \succ \{q\}$ and $\lambda^{\{p,q\}}(p) > 0$ then $\nu_u > 0$. Therefore $\sum_{s \in S, \alpha_s > -1} \eta_s > 0$, and by equation A.6, $V(\{p,q\}) > V(\{q\})$, which proves that $(\succeq, \lambda)$ satisfies axiom 5.1.

In addition, since $u(p) + \alpha_s(u(p) - u(q)) < u(p)$ when $\alpha_s < 0$, equation A.6 yields

$$\begin{aligned} V(\{p,q\}) &\leq \sum_{s \in S, \alpha_s > -1} \eta_s u(p) + \sum_{s \in S, \alpha_s < -1} \eta_s u(q) \\ &\leq \nu_u u(p) + \nu_{-u} u(q) = V(\{\lambda^{\{p,q\}}(p) p + \lambda^{\{p,q\}}(q) q\}) \end{aligned}$$

which proves that axiom 3.2 holds.

Proof of sufficiency Suppose first that $\nu_u = 0$. In that case, for any p, q such that $\{p\} \succ \{q\}$ we have $\lambda^{\{p,q\}}(p) = 0$, which by axiom 3.2 implies $\{p, q\} \sim \{q\}$. This identifies the Random Strotz representation associated with $\succeq$, given by $V = \min u$. This function attaches a weight equal to 1 to the state $-u$, therefore the joint representation is sophisticated.

If $\nu_u > 0$, let us define $\eta_u = \sum_{s \in S, \alpha_s > 0} \eta_s + \sum_{s \in S, 0 < \alpha_s < 1} \eta_s \alpha_s$ and rewrite equation A.6 under the form

$$V(x) = \eta_u \max_{p \in x} u(p) + (1 - \eta_u) \min_{q \in x} u(q)$$

Axiom 5.1 yields $\eta_u > 0$. If axiom 3.1 is satisfied, a direct adaptation of theorem 3.1 proves the existence of a sophisticated joint representation under the Random Strotz interpretation. If axiom 3.1 is violated, we obtain $\eta_u < \nu_u$. It is therefore possible to rewrite

$$V(x) = \nu_u [\max_{p \in x} \frac{\eta_u}{\nu_u} u(p) - \max_{q \in x} (\frac{\eta_u}{\nu_u} - 1) u(q)] + \nu_{-u} \min_{q \in x} u(q) \tag{A.7}$$

Equation A.7 defines a sophisticated representation of $\succeq$, since the function V attaches weights equal to (ν_u, ν_{-u}) to the states $(u, -u)$ respectively.

Appendix B: One-dimensional Random Strotz models

This appendix proves theorem 4.2, and provides another characterization of one-dimensional Random Strotz representations. Unlike the rest of the paper, this part does not assume any conditions on μ except non-triviality: continuity and finiteness are not required. The preference $\succeq$ is here defined over compact menus, as in Dekel and Lipman (2012).

1 Proof of Theorem 4.2

1.1 Necessity of axiom 4.4

Lemma B.4. *Suppose that $\succeq$ has a one-dimensional Random Strotz representation (u,μ) with direction $v \in \mathcal{V}$ such that $\mu(\mathcal{C}_v) > 0$. Consider two lotteries q_1 and q_2 verifying $\{q_1\} \sim \{q_2\}$, and suppose that there exists a lottery z such that $\{z\} \succ \{q_1\} \sim \{q_2\}$. The two following statements are equivalent: (i) There exists a menu x such that $\{p\} \succ \{q_1\} \sim \{q_2\}$ for any $p \in x$ and $x \cup \{q_1\} \succ x \cup \{q_2\}$; (ii) $v(q_1) < v(q_2)$.*

Proof. Let us prove $(i) \Rightarrow (ii)$ by contrapositive. $\{q_1\} \sim \{q_2\}$ implies $u(q_1) = u(q_2)$. Suppose that $v(q_1) \geq v(q_2)$, and take any x such that $u(p) > u(q_1)$. If $w \in \overline{\mathcal{C}}_v$, we can write $w = au + \sqrt{1-a^2}v$, where $|a| \leq 1$, and obtain $w(q_1) = au(q_1) + \sqrt{1-a^2}v(q_1) \geq au(q_2) + \sqrt{1-a^2}v(q_2) = w(q_2)$. Since the support of μ is included in $\overline{\mathcal{C}}_v$, q_1 dominates q_2 on all the possible subjective states. For $i = 1,2$, we write $\Omega^{x\cup\{q_i\}}(q_i) = \{w \in \overline{\mathcal{C}_v} | w(q_i) > \max_{p\in x} w(p)\}$ the list of subjective states on which q_i is chosen. The observation above yields $\Omega^{x\cup\{q_2\}}(q_2) \subseteq \Omega^{x\cup\{q_1\}}(q_1)$, and hence

$$
\begin{aligned}
V(x\cup\{q_1\}) &= \mu(\Omega^{x\cup\{q_1\}}(q_1))u(q_1) + \int_{w\notin\Omega^{x\cup\{q_1\}}(q_1)} \max_{p\in\mathcal{M}_w(x)} u(p)\mu(dw) \\
&\leq \mu(\Omega^{x\cup\{q_2\}}(q_2))u(q_2) + \int_{w\notin\Omega^{x\cup\{q_2\}}(q_2)} \max_{p\in\mathcal{M}_w(x)} u(p)\mu(dw) \\
&\leq V(x\cup\{q_2\})
\end{aligned}
$$

This proves that $x \cup \{q_2\} \succeq x \cup \{q_1\}$.

$(ii) \Rightarrow (i)$. Suppose that $u(q_1) = u(q_2)$ and $v(q_1) < v(q_2)$. Consider an increasing sequence $0 < a_n < 1$ of limit 1, and the increasing sequence of sets $\mathcal{C}_v(-a_n, a_n)$. This sequence has limit $\mathcal{C}_v$ which has positive measure, therefore by theorem 10.2 of Billingsley (2012) we have $\mu(\mathcal{C}_v(-a_n,a_n)) > 0$ for n large enough.

Define $a = a_n$.

Consider now a number γ such that: $0 < \gamma < \frac{\sqrt{1-a^2}}{2a}[v(q_2) - v(q_1)]$. Such a number exists since $a \in (0,1)$ and $v(q_2) > v(q_1)$. Suppose first that we can find γ small enough to ensure that $p = \frac{1}{2}q_1 + \frac{1}{2}q_2 + \gamma u$ is a legitimate lottery. Notice that $u(p) > u(q_1), u(q_2)$ and that

$$\begin{aligned} au(p) + \sqrt{1-a^2}v(p) &= a\gamma + au(q_2) + \sqrt{1-a^2}[\frac{1}{2}v(q_1) + \frac{1}{2}v(q_2)] \\ &< au(q_2) + \sqrt{1-a^2}v(q_2) \end{aligned}$$

Therefore q_2 dominates p over the set $\overline{\mathcal{C}}_v(-1, a)$. Similarly,

$$\begin{aligned} -au(p) + \sqrt{1-a^2}v(p) &= -a\gamma - au(q_1) + \sqrt{1-a^2}[\frac{1}{2}v(q_1) + \frac{1}{2}v(q_2)] \\ &> -au(q_1) + \sqrt{1-a^2}v(q_1) \end{aligned}$$

And hence p dominates q_1 over the set $\overline{\mathcal{C}}_v(-a, 1)$.

Thus we obtain $\Omega^{\{p,q_1\}}(q_1) \subset \Omega^{\{p,q_2\}}(q_2)$, and $\mathcal{C}_v(-a, a) \subseteq \Omega^{\{p,q_2\}}(q_2)\backslash\Omega^{\{p,q_1\}}(q_1)$. Since $\mu(\mathcal{C}_v(-a, a)) > 0$, this yields $\mu(\Omega^{\{p,q_1\}}(q_1)) < \mu(\Omega^{\{p,q_2\}}(q_2))$.

Hence

$$\begin{aligned} V(\{p, q_1\}) &= \mu(\Omega^{\{p,q_1\}}(q_1))u(q_1) + (1 - \mu(\Omega^{\{p,q_1\}}(q_1)))u(p) \\ &> \mu(\Omega^{\{p,q_2\}}(q_2))u(q_2) + (1 - \mu(\Omega^{\{p,q_2\}}(q_2)))u(p) = V(\{p, q_2\}) \end{aligned}$$

Therefore the triple $(x = \{p\}, q_1, q_2)$ satisfies $x \cup \{q_1\} \succ x \cup \{q_2\}$.

If $\dfrac{1}{2}q_1 + \dfrac{1}{2}q_2 + \gamma u$ is not a lottery for any $\gamma > 0$, since $\dfrac{1}{2}q_1 + \dfrac{1}{2}q_2$ is not a maximizer of u among $\Delta(\mathcal{Z})$, a standard separation argument shows that it is possible to find $\tilde{u} \in \mathcal{W}$ that satisfies $u.\tilde{u} > 0$ and such that $\dfrac{1}{2}q_1 + \dfrac{1}{2}q_2 + \gamma\tilde{u}$ is a lottery. The same construction holds with a large enough to satisfy $au.\tilde{u} > \sqrt{1-a^2}$ and γ such that $0 < \gamma[au.\tilde{u} \pm \sqrt{1-a^2}] < \dfrac{\sqrt{1-a^2}}{2}[v(q_2) - v(q_1)]$.

□

To prove the necessity of axiom 4.4, consider a one-dimensional Random Strotz (u, μ) with direction v. If $\mu(\mathcal{C}_v) = 0$, then $\mu(\{u, -u\}) = 1$, thus for any (x, q_1, q_2) such that $\{p\} \succ \{q_1\} \sim \{q_2\}$ for all $p \in x$, we have $x \cup \{q_1\} \sim x \cup \{q_2\}$, and the condition of axiom 4.4 is trivially satisfied.

Suppose now that $\mu(\mathcal{C}_v) > 0$ and take (x, y, q_1, q_2) such that $\{p\} \succ \{q_1\} \sim \{q_2\}$ for all $p \in x$ and $x \cup \{q_1\} \succ x \cup \{q_2\}$. Since q_1 and q_2 do not maximize u on $\Delta(\mathcal{Z})$, the implication $(i) \Rightarrow (ii)$ of lemma B.4 yields $v(q_1) < v(q_2)$, which in turn implies

$y \cup \{q_1\} \succeq y \cup \{q_2\}$. This completes the proof of necessity.

1.2 Sufficiency of axiom 4.4

Suppose that the representation (u, μ) of $\succeq$ is not one-dimensional. It is clear that $\mu(\mathcal{C}(-1,1)) > 0$, otherwise we would have $\mu(\{u, -u\}) = 1$ and therefore $\mu(\overline{\mathcal{C}}_v) = 1$ for any $v \in \mathcal{V}$.

Claim B.1. There exists $\xi < 1$ such that $\mu(\mathcal{C}_v) < \mu(\mathcal{C}(-\xi, \xi))$ for all $v \in \mathcal{V}$.

Proof. If this is not the case, there exists an increasing sequence $0 < a_n < 1$ of limit 1, and a sequence $v_n \in \mathcal{V}$ such that

$$\mu(\mathcal{C}_{v_n}) \geq \mu(\mathcal{C}(-a_n, a_n)) \tag{B.1}$$

Take m sufficiently large to guarantee $\mu(\mathcal{C}(-a_m, a_m)) > \frac{1}{2}\mu(\mathcal{C}(-1,1))$. This is possible since $\mathcal{C}(-a_m, a_m)$ is an increasing sequence of limit $\mathcal{C}(-1,1)$, which has positive measure. Suppose that v_n is not constant for $n \geq m$. We can find $v_1 \neq v_2$ such that $\mu(\mathcal{C}_{v_1}) > \frac{1}{2}\mu(\mathcal{C}(-1,1)$ and $\mu(\mathcal{C}_{v_2}) > \frac{1}{2}\mu(\mathcal{C}(-1,1)$, which implies $\mu(\mathcal{C}_{v_1} \cup \mathcal{C}_{v_2}) > \mu(\mathcal{C}(-1,1))$. This is a contradiction, since $\mathcal{C}_{v_1} \cup \mathcal{C}_{v_2} \subseteq \mathcal{C}(-1,1)$. Hence, v_n is constant for n large enough. Denote v^* its limit, and take the limit in B.1. We obtain $\mu(\mathcal{C}_v) \geq \mu(\mathcal{C}(-1,1))$, which further implies

$$\begin{aligned}\mu(\overline{\mathcal{C}}_v) &= \mu(\{u, -u\}) + \mu(\mathcal{C}_v) \\ &\geq \mu(\{u, -u\}) + \mu(\mathcal{C}(-1,1)) = 1\end{aligned}$$

And hence, $\mu(\overline{\mathcal{C}}_v) = 1$. This is a contradiction, since (u, μ) is not one-dimensional.

□

For each $v \in \mathcal{V}$, we define $\mathcal{B}(v, \epsilon) = \{w \in \mathcal{V} | \|w - v\| < \epsilon\}$ the open ball of radius ϵ and center v, restricted to utilities which are orthogonal to u. We also define

$$\mathcal{A}(v, \epsilon) = \bigcup_{w \in \mathcal{B}(v,\epsilon)} \mathcal{C}_w = \{au + \sqrt{1-a^2}w | -1 < a < 1, w \in \mathcal{V}, \|w - v\| < \epsilon\}$$

$\mathcal{A}(v, \epsilon)$ contains the utilities whose direction lies in the open ball of center v and radius ϵ.

Claim B.2. For any $\epsilon > 0$ low enough, there exists two expected utilities v_1^ϵ and v_2^ϵ such that $\mu(\mathcal{A}(v_1^\epsilon, \epsilon)) > 0$, $\mu(\mathcal{A}(v_2^\epsilon, \epsilon)) > 0$ and $\|v_1^\epsilon - v_2^\epsilon\| > 7\epsilon$.

Proof. Consider ξ defined by claim B.1. ξ can be chosen high enough to guarantee that $\mu(\overline{\mathcal{C}}(-\xi,\xi)) > 0$. The set $\overline{\mathcal{C}}(-\xi,\xi) = \{au + \sqrt{1-a^2}w, w \in \mathcal{V}, |a| \leq \xi\}$ is compact since it is closed and bounded. Moreover, it is covered by the union of the open sets $\mathcal{A}(v,\epsilon)$ for all $v \in \mathcal{V}$. By the Borel-Lebesgue theorem, there exists a finite family $\mathcal{F} \subset \mathcal{V}$ such that $\{\mathcal{A}(v,\epsilon)\}_{v\in\mathcal{F}}$ covers $\overline{\mathcal{C}}(-\xi,\xi)$. An immediate implication is that $\sum_{v\in\mathcal{F}} \mu(\mathcal{A}(v,\epsilon)) > 0$.

Let us show that, if ϵ is low enough, this subcovering contains at least two sets of positive measure that can be separated as stated in the claim. Let us proceed by contradiction. Suppose that for any $\eta > 0$, there exists $\epsilon < \eta$, a finite subset $\mathcal{F}_\epsilon$ such that $\{\mathcal{A}(v,\epsilon)\}_{v\in\mathcal{F}}$ covers $\overline{\mathcal{C}}(-\xi,\xi)$, and $v_\epsilon \in \mathcal{F}_\epsilon$ such that for any $v \in \mathcal{F}_\epsilon$, $\mu(\mathcal{A}(v,\epsilon)) > 0 \Rightarrow \|v - v_\epsilon\| < 7\epsilon$. Notice that $\bigcup_{v\in\mathcal{B}(v_\epsilon,7\epsilon)} \mathcal{A}(v,\epsilon) \subseteq \mathcal{A}(v_\epsilon, 8\epsilon)$. Hence, since $\sum_{v\in\mathcal{F}_\epsilon\cap\mathcal{B}(v_\epsilon,7\epsilon)} \mu(\mathcal{A}(v,\epsilon)) \geq \mu(\mathcal{C}(-\xi,\xi))$, we obtain $\mu(\mathcal{A}(v_\epsilon,8\epsilon)) \geq \mu(\mathcal{C}(-\xi,\xi))$.

Consider a decreasing sequence $\epsilon_n \to 0$. By the Bolzano-Weierstrass theorem, the sequence $v_n = v_{\epsilon_n}$ defined over the compact $\mathcal{V}$ admits a convergent subsequence. To simplify the notation, let us assume that v_n is itself convergent to a value v^*. For any n, we have $\mu(\mathcal{A}(v_n, 8\epsilon_n)) \geq \mu(\mathcal{C}(-\xi,\xi))$. The next step is to show that we can take the limit in this inequality and obtain $\mu(\mathcal{C}_{v^*}) \geq \mu(\mathcal{C}(-\xi,\xi))$.

Notice that

$$\mu(\mathcal{A}(v^*,8\epsilon_n)) - \mu(\mathcal{A}(v_n,8\epsilon_n)) = \mu(\bigcup_{w\in\mathcal{B}(v^*,8\epsilon_n)\backslash\mathcal{B}(v_n,8\epsilon_n)} \mathcal{C}_w) - \mu(\bigcup_{w\in\mathcal{B}(v_n,8\epsilon_n)\backslash\mathcal{B}(v^*,8\epsilon_n)} \mathcal{C}_w) \tag{B.2}$$

Consider the sets

$$\mathcal{G}_n = \bigcup_{m=n}^{+\infty} \bigcup_{w\in\mathcal{B}(v^*,8\epsilon_m)\backslash\mathcal{B}(v_m,8\epsilon_m)} \mathcal{C}_w$$

$\mathcal{G}_m$ is decreasing and has for limit $\lim_{n\to+\infty} \mathcal{G}_n = \emptyset$. Thus, $\lim_{n\to+\infty} \mu(\mathcal{G}_n) = 0$, and since $\bigcup_{w\in\mathcal{B}(v^*,8\epsilon_n)\backslash\mathcal{B}(v_n,8\epsilon_n)} \mathcal{C}_w \subseteq \mathcal{G}_n$, we obtain

$$\lim_{n\to+\infty} \mu(\bigcup_{w\in\mathcal{B}(v^*,8\epsilon_n)\backslash\mathcal{B}(v_n,8\epsilon_n)} \mathcal{C}_w) = 0$$

A similar argument proves that

$$\lim_{n\to+\infty} \mu(\bigcup_{w\in\mathcal{B}(v_n,8\epsilon_n)\backslash\mathcal{B}(v^*,8\epsilon_n)} \mathcal{C}_w) = 0$$

Moreover, the sequence of sets $\{\mathcal{A}(v^*, 8\epsilon_n)\}$ is decreasing and converges to $\mathcal{C}_{v^*}$

when $n \to +\infty$. Thus, taking the limit in equation B.2 shows that $\mu(\mathcal{A}(v_n, 8\epsilon_n)) \to \mu(\mathcal{C}_{v^*})$, which implies $\mu(\mathcal{C}_{v^*}) \geq \mu(\mathcal{C}(-\xi, \xi))$. This latter inequality contradicts the statement in claim B.1. This completes the proof of claim B.2. □

If $v \in \mathcal{V}$, we write $\mathcal{A}_v^\epsilon(-1, a^*) = \{au + \sqrt{1-a^2}w | -1 < a < a^*, w \in \mathcal{V} \cap \mathcal{B}(v, \epsilon)\}$: it contains the utilities whose direction lies in the open ball of center v and radius ϵ, and whose intensity lies strictly between -1 and by a^*.

Claim B.3. For any $\epsilon > 0$ small enough, there exists $v_1, v_2 \in \mathcal{V}$ and $a^* < 1$ such that
$\|v_1 - v_2\| > 7\epsilon$ and

$$\min(\mu(\mathcal{A}_{v_1}^\epsilon(-1, a^*)), \mu(\mathcal{A}_{v_2}^\epsilon(-1, a^*))) > \mu(\mathcal{C}(-1, -a^*))$$

Proof. Take v_1 and v_2 given by claim B.2. The proof stems immediately from the fact that $\mu(\mathcal{A}_{v_1}^\epsilon(-1, a^*))$ and $\mu(\mathcal{A}_{v_2}^\epsilon(-1, a^*))$ converge to positive values when $a^* \to 1$, while $\mu(\mathcal{C}(-1, -a^*))$ tends to zero. □

Claim B.4. For any ϵ small enough, there exists three numbers α, β, γ such that $\alpha > 0$,

$$\begin{cases} \alpha > \beta \\ \alpha \sup_{w \in \mathcal{B}(v_2, 3\epsilon)} w.v_1 < \beta \inf_{w \in \mathcal{B}(v_2, 3\epsilon)} w.v_2 \\ \alpha \sup_{w \in \mathcal{B}(v_1, 3\epsilon)} w.v_2 < \beta \inf_{w \in \mathcal{B}(v_1, 3\epsilon)} w.v_1 \end{cases}$$

and

$$\begin{cases} \alpha < \gamma \\ \alpha \inf_{w \in \mathcal{B}(v_1, \epsilon)} w.v_1 > \gamma \sup_{w \in \mathcal{B}(v_1, \epsilon)} w.v_2 \\ \alpha \inf_{w \in \mathcal{B}(v_2, \epsilon)} w.v_2 > \gamma \sup_{w \in \mathcal{B}(v_2, \epsilon)} w.v_1 \end{cases}$$

Proof. Suppose that $w \in \mathcal{B}(v_1, \epsilon)$. We have

$$\begin{aligned} w.v_1 &= \frac{1}{2}(\|w\|^2 + \|v_1\|^2 - \|w - v_1\|^2) \\ &= 1 - \frac{1}{2}\|w - v_1\|^2 \\ &> 1 - \frac{1}{2}\epsilon^2 \end{aligned}$$

And since $\|w-v_2\| \geq \|v_1-v_2\| - \|w-v_1\| \geq 6\epsilon$, we also obtain $w.v_2 < 1 - \frac{1}{2}(6\epsilon)^2$. Similarly, $\inf_{w\in\mathcal{B}(v_2,\epsilon)} w.v_2 > 1 - \frac{1}{2}\epsilon^2$, and $\sup_{w\in\mathcal{B}(v_2,\epsilon)} w.v_1 < 1 - \frac{1}{2}(6\epsilon)^2$. Hence, if ϵ is small enough for the denominators to be positive, the ratios $\frac{\sup_{w\in\mathcal{B}(v_2,\epsilon)} w.v_1}{\inf_{w\in\mathcal{B}(v_2,\epsilon)} v.w_2}$ and $\frac{\sup_{w\in\mathcal{B}(v_1,\epsilon)} w.v_2}{\inf_{w\in\mathcal{B}(v_1,\epsilon)} w.v_1}$ are bounded above away from 1. It is thus easy to find α and γ that satisfy the conditions. A similar reasoning can be applied to find β. □

To complete the proof, take $\phi > 0$, ϵ low enough to ensure that $w.v_1 > 0$ as soon as $w_1 \in \mathcal{B}(v_1, 3\epsilon)$, and $w.v_2 > 0$ as soon as $w_2 \in \mathcal{B}(v_2, 3\epsilon)$, and take $a^*, v_1, v_2, \alpha, \beta, \gamma$ defined by claims B.3 and B.4. Define $p_1 = \mathbb{1} + \epsilon(\phi u + \gamma v_1)$, $p_2 = \mathbb{1} + \epsilon(\phi u + \beta v_2)$, $r_1 = \mathbb{1} + \epsilon(\phi u + \beta v_1)$ and $r_2 = \mathbb{1} + \epsilon(\phi u + \gamma v_2)$, and for any $w \in \mathcal{V}$ such that $w \notin \mathcal{B}(v_1, 3\epsilon) \cup \mathcal{B}(v_2, 3\epsilon)$, $z_w = \mathbb{1} + \epsilon(\phi u + \alpha w)$. u is indifferent between all these elements, with valuation $\epsilon\phi$. Consider the menu x containing p_1, p_2 and all the z_w, and the menu y containing r_1, r_2 and all the z_w.

Define now $q_1 = \mathbb{1} + \epsilon\alpha v_1$ and $q_2 = \mathbb{1} + \epsilon\alpha v_2$. We have $u(q_1) = u(q_2) = 0 < \epsilon\phi$, hence q_1 and q_2 are normatively inferior to all elements of x and y. The next step is to show that q_1 is more tempting than q_2 with respect to the menu y, while q_2 is more tempting than q_1 with respect to the menu y.

Consider the *ex post* choice in $x \cup \{q_1\}$. It is clear that q_1 is chosen by $-u$ and not chosen by u. Consider $w \notin \{u, -u\}$, and write $w = au + \sqrt{1-a^2}\overline{w}$, where $-a^* \leq a \leq 1$. We have

$$w(q_1) = \epsilon\sqrt{1-a^2}\alpha\overline{w}.v_1 \tag{B.3}$$

Suppose first, that $\overline{w} \in \mathcal{B}(v_1, 3\epsilon)$. We have

$$w(p_1) = a\epsilon\phi + \epsilon\sqrt{1-a^2}\gamma\overline{w}.v_1 \tag{B.4}$$

Compare B.3 with B.4. Since $\gamma > \alpha$ and $a^* < 1$ one can pick ϕ low enough to impose the inequality $w(p_1) > w(q_1)$ for all values of $a \geq a^*$. Given this choice, p_1 dominates q_1 if $\overline{w} \in \mathcal{B}(v_1, 3\epsilon)$ and $a \geq -a^*$.

Suppose now that $\overline{w} \in \mathcal{B}(v_2, 3\epsilon)$. We have

$$w(p_2) = a\epsilon\phi + \epsilon\sqrt{1-a^2}\beta\overline{w}.v_2 \tag{B.5}$$

Compare B.3 with B.5, and notice that $\beta \inf \overline{w}.v_2 > \alpha \sup \overline{w}.v_1$ by claim B.4. Hence, again, if ϕ is low enough, p_2 dominates q_1 if $\overline{w} \in \mathcal{B}(v_2, 3\epsilon)$ and $a \geq -a^*$.

Finally, suppose that $\overline{w} \notin \mathcal{B}(v_1, 3\epsilon) \cup \mathcal{B}(v_2, 3\epsilon)$. Notice that

$$w(z_{\overline{w}}) = a\epsilon\phi + \epsilon\sqrt{1-a^2}\alpha \tag{B.6}$$

Compare B.3 and B.6. $v_1.\overline{w}$ is uniformly bounded away from 1 since $\|\overline{w} - v_1\| \geq 3\epsilon$. Hence, by the same argument as above, if ϕ is small enough, q_1 is dominated by $z_{\overline{w}}$ if $\overline{w} \notin \mathcal{B}(v_1, 3\epsilon) \cup \mathcal{B}(v_2, 3\epsilon)$ and $a \geq -a^*$.

To sum up, if ϕ is low enough, q_1 is never chosen in $x \cup \{q_1\}$ by a state whose intensity of temptation is stronger than $-a^*$. Therefore, denoting $\alpha^{x\cup\{q_1\}} = \mu(\{w : w(q_1) > \max_{p\in x} w(p)\})$ the *ex ante* probability of choosing q_1 we have

$$\alpha^{x\cup\{q_1\}}(q_1) \leq \mu(\{-u\}) + \mu(\mathcal{C}(-1, a^*)) \tag{B.7}$$

Consider now the anticipated choice in the set $y \cup \{q_1\}$ in a state of the form $w = au + \sqrt{1-a^2}\overline{w}$ where $a < a^*$ and $\overline{w} \in \mathcal{B}(v_1, \epsilon)$. Notice that

$$w(r_1) = a\epsilon\phi + \epsilon\sqrt{1-a^2}\beta\overline{w}.v_1 \tag{B.8}$$

Since $\alpha > \beta$, $v_1.\overline{w} > 0$ and $a < a^* < 1$, by equations B.3and B.8 it is possible to take ϕ low enough to ensure that $w(r_1) < w(q_1)$. Observe now that

$$w(r_2) = a\epsilon\phi + \epsilon\sqrt{1-a^2}\gamma\overline{w}.v_2 \tag{B.9}$$

But by claim B.4, we have $\alpha \inf \overline{w}.v_1 > \gamma \sup \overline{w}.v_2$. Hence we can choose ϕ low enough to ensure that $w(r_2) < w(q_1)$.

Finally, for any $\hat{w} \notin \mathcal{B}(v_1, 3\epsilon) \cup \mathcal{B}(v_2, 3\epsilon)$ it is easy to show that $\hat{w}.\overline{w} < \overline{w}.v_1$. Thus, since $w(z_{\hat{w}}) = a\epsilon\phi + \epsilon\alpha\sqrt{1-a^2}\hat{w}.\overline{w}$, we obtain $w(z_{\hat{w}}) < w(q_1)$ for any $\hat{w}$ as soon as $a < a^*$, if ϕ is chosen small enough.

To sum up, q_1 is the single maximizer $y \cup \{q_1\}$ at least on the utilities of the form $au + \sqrt{1-a^2}\overline{w}$, where $\overline{w} \in \mathcal{B}(v_1, \epsilon)$ and $a < a^*$, which means that the probability of choosing it verifies

$$\alpha^{y\cup\{q_1\}}(q_1) \geq \mu(\{-u\}) + \mu(\mathcal{A}^{\epsilon}_{v_1}(-1, a^*)) \tag{B.10}$$

The same arguments prove that the choice of q_2 in x or y satisfies

$$\alpha^{x\cup\{q_2\}}(q_2) \geq \mu(\{-u\}) + \mu(\mathcal{A}^{\epsilon}_{v_2}(-1, a^*)) \tag{B.11}$$

and

$$\alpha^{y\cup\{q_2\}}(q_2) \leq \mu(\{-u\}) + \mu(\mathcal{C}(-1, -a^*)) \tag{B.12}$$

Compare equations B.7 and B.11. By claim B.3, we obtain $\alpha^{x\cup\{q_1\}}(q_1) < \alpha^{x\cup\{q_2\}}(q_2)$, which implies $x \cup \{q_1\} \succ x \cup \{q_2\}$. In contrast, by equations B.10 and B.12, $\alpha^{y\cup\{q_1\}}(q_1) > \alpha^{y\cup\{q_2\}}(q_2)$, which implies $y \cup \{q_1\} \prec y \cup \{q_2\}$. These two properties together violate axiom 4.4.

2 An alternative characterization

2.1 Representation theorem

This section shows that one-dimensional Random Strotz representations are characterized by another intuitive behavioral property.

Axiom B.1 (Unique Temptation).
If $\{p\} \succ \{q\}$ for any $p \in x\cup y$, then $x \succ x\cup\{q\}, y \succ y\cup\{q\} \Rightarrow x\cup y \succ x\cup y\cup\{q\}$.

Axiom B.1 also characterizes one-dimensional models, with one caveat: if the support of μ cannot be bounded away from $-u$, the consequent $x\cup y \succ x\cup y\cup\{q\}$ is always true because all inferior options are tempting. Axiom B.1 has no content in that case. To overcome this issue, it is possible to impose another property that guarantees the existence of a higher bound on the intensity of the temptation.

Axiom B.2 (Limited Temptation).
$\exists x \in \mathcal{A}, q \in \Delta(Z)$ such that $\{p\} \succ \{q\}$ for any $p \in x \cup y$ and $x \sim x \cup \{q\}$.

Axiom B.2 is a richness condition: it simply states that some options in the choice set are not tempting. This condition is innocuous, since it is satisfied if an option that appears extremely undesirable both *ex ante* and *ex post* is added to the set of prizes. The proof will first show that axiom B.2 is equivalent to the existence of a a neighborhood of $\{-u\}$ of measure zero; and then proceed to show that, among the Random Strotz models that satisfy axiom B.2, one-dimensional models are characterized by axiom B.1.

Lemma B.5. *Suppose that $\succeq$ has a Random Strotz representation (u, μ). The following statements are equivalent: (i) $\succeq$ satisfies axiom B.2; (ii) there exists $a > -1$ such that $\mu(\overline{\mathcal{C}}(-1, a)) = 0$.*

Proof. $(ii) \Rightarrow (i)$. Suppose that μ satisfies (ii) for some $a > -1$. Consider a pair (ϵ, γ) such that $\epsilon > 0$ and $a\epsilon + \sqrt{1-a^2}\gamma > 0$. Define $q = \mathbb{1}$ and $p_v = \mathbb{1} + \phi(\epsilon u + \gamma v)$

for $v \in \mathcal{V}$, where $\phi > 0$ is taken sufficiently small for p_v to be an interior lottery for all v. We have $u(p_v) = \phi\epsilon > u(q) = 0$ for all v, and

$$au(p_v) + \sqrt{1-a^2}v(p_v) = \phi(a\epsilon + \sqrt{1-a^2}\gamma) > 0 = au(q) + \sqrt{1-a^2}v(q)$$

This shows that $\alpha^{x\cup\{q\}}(q) \leq \mu(\overline{\mathcal{C}}(-1,a)) = 0$, and and hence $x \sim x \cup \{q\}$.

$(i) \Rightarrow (ii)$. Suppose that (ii) does not hold, i.e. that for any $a > -1$, we have $\mu(\overline{\mathcal{C}}(-1,a)) > 0$.

Consider now x and q such that $u(p) > u(q)$ for any $p \in x$. Consider the function $a : \mathcal{V} \times x \Rightarrow (-1,1)$ defined by the equation

$$\frac{a(v,p)}{\sqrt{1-a(v,p)^2}} = \frac{v(q)-v(p)}{u(p)-u(q)} \tag{B.13}$$

equation B.13 uniquely defines a value $a(v,p)$ such that $a(v,p) > -1$. Moreover, a is continuous in the product topology. By Tychonoff's theorem, $\mathcal{V} \times x$ is compact since $\mathcal{V}$ and x are compact. Thus, $\inf_{(v,p)\in\mathcal{V}\times x} a(v,p) > -1$. Take any a such that $-1 < a < \inf_{(v,p)\in\mathcal{V}\times x} a(v,p)$, any $(v,p) \in \mathcal{X} \times a$. By equation B.13 we have $au(p) + \sqrt{1-a^2}v(p) < au(q) + \sqrt{1-a^2}v(q)$. Hence, $\Omega^{x\cup\{q\}} \supseteq \overline{\mathcal{C}}(-1,a)$, which yields $\mu(\Omega^{x\cup\{q\}}) > 0$, and finally $x \cup \{q\} \prec x$. Since this result is obtained for any pair (x,q), the preference $\succeq$ does not satisfy axiom B.2. □

Theorem B.1. *Suppose that $\succeq$ has a Random Strotz representation (u,μ), and that $\succeq$ satisfies axiom B.2. (u,μ) is one-dimensional if and only if $\succeq$ satisfies axiom B.1.*

2.2 Necessity of axiom B.1

Suppose that $\succ$ has a one-dimensional Random Strotz representation (u,μ) of direction v. Take a triple (x,y,q) such that $\{p\} \succ \{q\}$ for any $p \in x \cup y$, $x \succ x \cup \{q\}$ and $y \succ y \cup \{q\}$. The same arguments used to find a in the proof of lemma B.5 can be used to obtain

$$a_x = \sup\{a \in [-1,1] | au(q) + \sqrt{1-a^2}v(q) \geq \sup_{p\in x} au(p) + \sqrt{1-a^2}v(p)\}$$

and a_y similarly by substituting y for x. $x \succ x \cup \{q\}$ and $y \succ y \cup \{q\}$ imply $\mu(\{-u\}) + \mu(\mathcal{C}_v(-1,a_x)) > 0$ and $\mu(\{-u\}) + \mu(\{-u\}) + \mu(\mathcal{C}_v(-1,a_y)) > 0$. Define $a = \min(a_x, a_y)$. It is easy to see that q is chosen in the menu $x \cup y \cup \{q\}$ by $-u$ and by all the utilities of the form $\tilde{a}u + \sqrt{1-\tilde{a}^2}v$ where $-1 < \tilde{a} < a$, and that this set has positive measure. As a consequence, $x \cup y \succ x \cup y \cup \{q\}$.

2.3 Sufficiency of axiom B.1

Let us prove the sufficiency of the axiom by contrapositive. Suppose that the Random Strotz representation (u, μ) of $\succ$ is not one-dimensional, and that $\succ$ satisfies axiom B.2.

Claim B.5. For any $\epsilon > 0$ low enough, there exists $v_1, v_2 \in \mathcal{V}$ and $a^* \in (0,1)$ such that: (i) $\mu(\mathcal{A}^{\epsilon}_{v_1}(-1, a^*)) > 0$, (ii) $\mu(\mathcal{A}^{\epsilon}_{v_2}(-1, a^*)) > 0$, (iii) $\|v_1 - v_2\| > 7\epsilon$, and (iv) $\mu(\overline{\mathcal{C}}(-1, -a^*))) = 0$.

Proof. The first three parts come from claims B.2 and B.3 in the proof of theorem 4.2. Part *(iv)* comes from lemma B.5. □

Take $\alpha, \phi > 0$ and define now $q = \mathbb{1}$, and: (i) for any $\hat{w} \notin \mathcal{B}(v_1, 3\epsilon)$, $p_{\hat{w}} = \mathbb{1} + \phi u + \alpha(\hat{w} - v_1)$; (ii) for any $\hat{w} \notin \mathcal{B}(v_2, 3\epsilon)$, $r_{\hat{w}} = \mathbb{1} + \phi u + \alpha(\hat{w} - v_2)$. α and ϕ can be taken small enough to make sure that these elements are well-defined lotteries. Define also $x = \{p_{\hat{w}}\}_{\hat{w} \notin \mathcal{B}(v_1, 3\epsilon)}$ and $y = \{r_{\hat{w}}\}_{\hat{w} \notin \mathcal{B}(v_2, 3\epsilon)}$. We observe that u equals $\phi > u(q) = 0$ on any element of $x \cup y$.

Consider the choice made in $x \cup \{q\}$. Take $w \in \mathcal{W}$, written $w = au + \sqrt{1-a^2}\,\overline{w}$. Suppose that $\overline{w} \in \mathcal{B}(v_1, \epsilon)$. We have $w(q) = 0$, and for any $\hat{w} \notin \mathcal{B}(v_1, 3\epsilon)$,

$$w(p_{\hat{w}}) = a\phi + \sqrt{1-a^2}\,\alpha(\hat{w}.\overline{w} - v_1.\overline{w})$$

In addition, we have

$$\begin{aligned} \hat{w}.\overline{w} &= \frac{1}{2}(\|\hat{w}\|^2 + \|\overline{w}\|^2 - \|\overline{w} - \hat{w}\|^2) \\ &= 1 - \frac{1}{2}\|\overline{w} - \hat{w}\|^2 \\ &\leq 1 - 2\epsilon^2 \end{aligned}$$

since $\|\overline{w} - \hat{w}\| \geq \|\hat{w} - v_1\| - \|\overline{w} - v_1\| \geq 2\epsilon$.

A similar argument shows that $v_1.\overline{w} \geq 1 - \frac{\epsilon^2}{2}$, which implies $\hat{w}.\overline{w} - v_1.\overline{w} < -\frac{3\epsilon^2}{2} < 0$. Therefore we can choose ϕ small enough such that the inequality $w(p_{\hat{w}}) < 0$ is satisfied provided that $a < a^*$. We obtain $\Omega^{x \cup \{q\}}(q) \supseteq \mathcal{A}^{\epsilon}_{v_1}(-1, a^*)$, which implies $\alpha^{x \cup \{q\}}(q) > 0$, and hence $x \cup \{q\} \prec x$. Similarly, we choose ϕ small enough to obtain $y \cup \{q\} \prec y$.

Suppose now that $\overline{w} \notin \mathcal{B}(v_1, 3\epsilon)$. We have

$$w(p_{\overline{w}}) = a\phi + \sqrt{1-a^2}\alpha(1 - v_1.\overline{w})$$

And since $\|v_1 - \overline{w}\| \geq 3\epsilon$, $v_1.\overline{w} < 1 - \frac{9\epsilon^2}{2}$. Therefore we can choose ϕ low enough to ensure that $w(p_{\overline{w}}) > 0$ is satisfied as soon as $a \geq -a^*$. Similarly, if ϕ is small enough and $\overline{w} \notin \mathcal{B}(v_2, 3\epsilon)$, the inequality $w(r_{\overline{w}}) > 0$ is satisfied if $a \geq -a^*$. Since $\|v_1 - v_2\| > 7\epsilon$, $(\mathcal{W} \setminus \mathcal{B}(v_1, 3\epsilon)) \cup (\mathcal{W} \setminus \mathcal{B}(v_2, 3\epsilon)) = \mathcal{W}$. This proves that, if $a \geq -a^*$, in every direction $\overline{w}$, q is dominated by an element of $x \cup y$. Therefore $\alpha^{x \cup y \cup \{q\}}(q) \leq \mu(\overline{\mathcal{C}}(-1, -a^*)) = 0$, and thus $x \cup y \cup \{q\} \sim x \cup y$. The triple (x, y, q) violates axiom B.1.

Appendix C: Experimental Instructions

1 Initial session

Page 1

Welcome! Thank you for participating in this decision-making experiment. Please read carefully the explanations below.

Rules You are invited to participate in an experiment on intertemporal decision-making. Participation is voluntary, and you are free to leave the experiment at any moment.

Risks and benefits There is no risk to participating in this study. You will be compensated for your time according to the quality of your answers.

Procedure The experiment will be run on your personal computer. Detailed instructions will be given on the next page.

Contact information If you have any questions about this research, you can contact the protocol director, Yves Le Yaouanq, through Mechanical Turk or at the following e-mail address: yves.le-yaouanq@m4x.org. This research is supported and approved by the Toulouse School of Economics, a European research institution.

Confidentiality No personal information will be recorded except your Mechanical Turk Worker ID, which allows us to proceed with the payment on the Mechanical Turk platform. This information will not be used otherwise and will be removed from our data as soon as the experiment is over.

If you consent to participating, please check the box below and click on "Submit" to go further.

I have read and understood these explanations, and I consent to participating in the experiment described. ○

Submit

Page 2

Participation fee Please read carefully these instructions. Some questions will be asked on the next page to verify that you read and understood these explanations. If your answers are correct, your participation will be validated and you will receive $1 (1 dollar) as a participation fee within the next 24 hours.

Additional earnings This experiment gives you the possibility to participate in 10 experimental sessions. You will be able to earn $0.40 (forty cents) for each session, which means up to $4 in total, in addition to the $1 participation fee.

Task The task is very simple. It consists in remembering to log in to the experimental website with your Mechanical Turk ID. The sessions will take place every day starting tomorrow during 10 days, from Sunday 31st of May to Tuesday 9th of June included.

At each of these dates, you will earn $0.40 if you log in to the website within the session, and $0 if you forget. The beginning and the end of the sessions are set on midnight according to the US Eastern Time Zone, the time zone of New York City. Your earnings will be paid as a bonus on the Mechanical Turk platform on Wednesday 10th of June.

You will not receive any reminder from us during the whole experiment. You are free to set a reminder yourself if you wish.

Website The URL of the experimental website is **********. You need to enter your Mechanical Turk ID to log in to the website. A confirmation message will be displayed and will inform you of your earnings so far and of the dates of the remaining sessions. Once you have carefully read and understood these instructions, please click on "Next" below.

Next

Page 3

Instructions Please read carefully the instructions below. Some questions will be asked at the bottom of this page. If you answer them correctly, you will receive $1 as a baseline participation fee within the next 24 hours; in addition, you will be allowed to participate in the future experimental sessions. If you provide any wrong or missing answer, your participation will not be validated and the experiment will be over immediately.

Payment rule To complete this initial session, we give you the possibility to modify the payment rule for one of the 10 experimental sessions. This session will be chosen randomly with equal probabilities among the 10 dates, and the payment rule for the 9 other dates will be unchanged. You will not learn today which date has been chosen, but you will be informed of this at the end of the experiment.

Principle For each of the following 21 rows, you are asked to indicate the option that you prefer among the two columns, left and right. The left column ("$0.40 if you log in") is similar to the rule used for the 9 other dates: if you choose this option, you will receive $0.40 for the session if you visit the website during the day, and $0 if you forget.

In contrast, if you choose the right column, you receive $0.40 with some probability, no matter if you visit the website or not. This option allows you to earn a chance of receiving $0.40 even if you forget to sign in. As you proceed down the rows, the chance of receiving $0.40 decreases. For instance, at row 1, the right column means that you receive $0.40 for sure. At row 10, it means that you receive $0.40 with probability 50%. At row 21, it means that you never receive $0.40. This payment is computed by the server and does not depend on whether you visit the website during the session or not.

For each of the 21 rows, you are asked to indicate your preference between the two columns. One of these rows will be selected randomly (with equal probabilities) and your choice will be recorded and implemented, so it is in your best interest to choose what you truly prefer between both options, for each row.

Remember that this choice will be used only for 1 of the 10 days, so it is in your best interest to try to remember the task and to log in everyday. In addition, recall that you won't receive any reminder from us during the experiment.

Your choice Please indicate your preferred option for each of the following 21 rows.

Left: you earn...	Right: you earn...
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 100% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 95% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 90% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 85% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 80% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 75% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 70% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 65% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 60% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 55% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 50% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 45% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 40% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 35% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 30% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 25% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 20% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 15% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 10% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 5% chance, \$0 otherwise
○ \$0.40 if you log in, \$0 otherwise	○ \$0.40 with 0% chance, \$0 otherwise

Final questions To verify that you read and understood the above instructions, you are now asked to answer two attention questions. Providing a correct answer to both questions is required to receive your \$1 participation fee and to be allowed to participate in the future sessions. Otherwise, the experiment will be over immediately and your participation will not be validated.

Suppose that row 3 is chosen, and that you chose the left column. If you remember to visit the website within the session, how much will you earn for this session?

○ $0.40 for sure ○ $0.40 with probability 90%

Suppose that row 3 is chosen, and that you chose the right column. If you forget to visit the website within the session, how much will you earn for this session?

○ $0.40 for sure ○ $0.40 with probability 90%

Submit

Page 4

Thank you for your answers. Your answers are correct. You are allowed to proceed with the experimental sessions. Your participation will be validated. Please copy and paste the following personal code in the Mechanical Turk online form to receive your $1 participation fee in the next 24 hours: *********. The initial experimental session is over. You can now close the window or log in to the website again to review the schedule of your future sessions by clicking on this *link*. Thank you for your participation.

---

2 Regular session (example)

Welcome! Today is Wednesday 3rd of June. There is an experimental session today. You will earn $0.40 for your participation today.

Your history 4 sessions have been run until now. You participated in 1 session so far, on Wednesday 3rd of June.

You missed 3 sessions so far, at the dates listed below: Sunday 31st of May Monday 1st of June Tuesday 2nd of June

Your schedule The future sessions will take place on: Thursday 4th of June Friday 5th of June Saturday 6th of June Sunday 7th of June Monday 8th of June Tuesday 9th of June

Remember that your total earnings will be paid on the Mechanical Turk platform on Wednesday 10th of June. Good luck!

References

Agranov, Marina and Pietro Ortoleva (forthcoming), "Stochastic choice and preferences for randomization." *Journal of Political Economy*.

Ahn, David S, Ryota Iijima, and Todd Sarver (2015), "Comparative measures of naiveté." Working paper.

Ahn, David S and Todd Sarver (2013), "Preference for flexibility and random choice." *Econometrica*, 81, 341–361.

Azrieli, Yaron, Christopher P Chambers, and Paul J Healy (2015), "Incentives in experiments: a theoretical analysis." Working paper.

Bartling, Björn, Ernst Fehr, and Holger Herz (2014), "The intrinsic value of decision rights." *Econometrica*, 82, 2005–2039.

Baumeister, Roy F, Kathleen D Vohs, and Dianne M Tice (2007), "The strength model of self-control." *Current directions in psychological science*, 16, 351–355.

Billingsley, Patrick (2012), *Probability and Measure*. Wiley.

Bisin, Alberto and Kyle Hyndman (2015), "Present-bias, procrastination and deadlines in a field experiment." Working paper.

Dean, Mark and John McNeill (2015), "Preference for flexibility and random choice: an experimental analysis." Working paper.

Dekel, Eddie and Barton Lipman (2012), "Costly self-control and random self-indulgence." *Econometrica*, 80, 1271–1302.

Dekel, Eddie, Barton Lipman, and Aldo Rustichini (2001), "Representing preferences with a unique subjective state space." *Econometrica*, 69, 891–934.

Della Vigna, Stefano and Ulrike Malmendier (2004), "Contract design and self-control: theory and evidence." *Quarterly Journal of Economics*, 119, 353–402.

Dwenger, Nadja, Dorothea Kübler, and Georg Weizsacker (2014), "Flipping a coin: Theory and evidence." Working paper.

Eliaz, Kfir and Ran Spiegler (2006), "Contracting with diversely naive agents." *Review of Economic Studies*, 73, 689–714.

Ergin, Haluk and Todd Sarver (2010), "A unique costly contemplation representation." *Econometrica*, 78, 1285–1339.

Ericson, Keith Marzilli (2011), "Forgetting we forget: overconfidence and prospective memory." *Journal of the European Economic Association*, 9, 43–60.

Exley, Christine and Jeffrey K Naecker (2015), "Observability increases the demand for commitment devices." Working paper.

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and self-control." *Econometrica*, 69, 1403–1435.

Gul, Faruk and Wolfgang Pesendorfer (2006), "Random expected utility." *Econometrica*, 74, 121–146.

Heidhues, Paul and Botond Kőszegi (2010), "Exploiting naivete about self-control in the credit market." *American Economic Review*, 100, 2279–2303.

Horton, John J, David G Rand, and Richard J Zeckhauser (2011), "The online laboratory: conducting experiments in a real labor market." *Experimental economics*, 14, 399–425.

Kőszegi, Botond (2014), "Behavioral contract theory." *Journal of Economic Literature*, 52, 1075–1118.

Kopylov, Igor (2012), "Perfectionism and choice." *Econometrica*, 80, 1819–1843.

Kreps, David M (1979), "A preference for flexibility." *Econometrica*, 47, 565–576.

Laibson, David (1997), "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 112, 443–478.

Lipman, Bart and Wolfgang Pesendorfer (2013), "Temptation." In *Advances in Economics and Econometrics: Tenth World Congress, volume 1* (Daron Acemoglu, Manuel Arellano, and Eddie Dekel, eds.), Cambridge University Press.

Mahajan, Aprajit and Alessandro Tarozzi (2011), "Time inconsistency, expectations and technology adoption: The case of insecticide treated nets." Working paper.

Mason, Winter and Siddharth Suri (2012), "Conducting behavioral research on amazon's mechanical turk." *Behavioral Research Methods*, 44, 1–23.

Noor, Jawwad (2007), "Commitment and self-control." *Journal of Economic Theory*, 135, 1–34.

Olea, Jose Luis Montiel and Tomasz Strzalecki (2014), "Axiomatization and measurement of quasi-hyperbolic discounting." *Quarterly Journal of Economics*, 129, 1449–1499.

Ortoleva, Pietro (2013), "The price of flexibility: towards a theory of thinking aversion." *Journal of Economic Theory*, 148, 903–934.

Owens, David, Zachary Grossman, and Ryan Fackler (2014), "The control premium: a preference for payoff autonomy." *American Economic Journal: Microeconomics*, 6, 138–161.

Rabin, Matthew and Ted O'Donoghue (1999), "Doing it now or later." *American Economic Review*, 89, 103–124.

Stovall, John (2010), "Multiple temptations." *Econometrica*, 78, 349–376.

Stovall, John (2015), "Temptation with uncertain normative preferences." Working paper.

Strotz, Robert Henry (1956), "Myopia and inconsistency in dynamic utility maximization." *Review of Economic Studies*, 23, 165–180.

Tversky, Amos (1969), "Intransitivity of preferences." *Psychological review*, 76, 31.