IDENTIFYING BELIEF-DEPENDENT PREFERENCES

ENRICO MATTIA SALONIA*

Job Market Paper

Latest version here

November 10, 2024

Abstract

Why are investors overconfident and trade excessively? Why do patients at health risk avoid testing? Why do voters choose different information sources? Possibly because their beliefs directly influence their well-being, i.e., they have belief-dependent preferences. However, existing theories of belief-dependent preferences struggle to generate testable predictions or to identify simultaneously beliefs and preferences. This paper addresses these issues by providing an axiomatic characterization of a class of preferences and belief-updating rules that deviate from Bayesian updating. Preferences, beliefs, and updating rules are identified from choices over contingent menus, each entailing a menu of acts available at a later time contingent on an uncertain state of the world. The results provide a theory-based approach to experimental designs to test information avoidance, distortion, and other behaviours consistent with belief-dependent preferences.

*Toulouse School of Economics. mattia.salonia@tse-fr.eu. I am indebted to Ingela Alger, François Salanié and Marciano Siniscalchi. I would also like to thank David Austen-Smith, Lony Bessagnet, Milo Bianchi, Michele Bisceglia, Annalisa Costella, Johannes Hörner, Alexander Jakobsen, Pau Juan-Bartroli, Peter Klibanoff, Moritz Loewenfeld, Chiara Margaria, Sophie Moinas, Maximilian Müller, Pedram Pooyafar, Sébastien Pouget, Philip Reny, Ernesto Rivera Mora, Yuval Salant, Sara Shahanaghi, Karine Van der Straeten, worshop participants at TSE, Northwestern and Autònoma Barcelona, and conference participants at the CEA 2024 in Toronto, NASMES 2024 in Nashville and SAET-EWET 2024 in Manchester. I acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 789111 - ERC EvolvingEconomics).

1 INTRODUCTION

People often hold some beliefs dear, even when faced with evidence against them. Do they change their views, or do they continue believing what they want? Indeed, research documents two common attitudes towards new information: avoidance and distortion. Investors avoid obtaining information when they expect the market to be in a bad state (Karlsson et al., 2009; Sicherman et al., 2016). Donors do not learn about the impact of their donation (Andreoni et al., 2017; Chan et al., 2024). Voters shun evidence that undermines their view (Gentzkow & Shapiro, 2010; Bakshy et al., 2015). Despite being at risk, patients refrain from getting tested. (Thornton, 2008; Oster et al., 2013). If information is obtained, it is distorted when unwelcome but correctly processed if welcomed. Negotiators interpret favourably information that supports their case and dismiss the quality of contrary evidence (Babcock et al., 1995). Individuals are less likely to accept bad news about their attractiveness or intelligence but correctly process good news (Eil & Rao, 2011). They also exaggerate good news about their ability or IQ (Möbius et al., 2022; Drobner & Goerg, 2024). Those with differing opinions on climate change consider scientists experts only when they support their views (Kahan et al., 2012).[1]

These observations conflict with expected utility theory, where more information is always desirable, and information should be processed via Bayes rule. However, if an individual derives pleasure or pain from specific beliefs, she might avoid information or process it differently. Theories of belief-dependent preferences (hereafter BDP) assume that individuals hold beliefs to enhance their well-being, which explains why they might avoid or distort information. Examples include psychological expected utility (Caplin & Leahy, 2001), optimal expectations (Brunnermeier & Parker, 2005), ego concerns (Köszegi, 2006), preferences for anticipation (Köszegi, 2010), and motivated reasoning (Bénabou, 2015; Bénabou & Tirole, 2016). These theories share a common feature: they do not provide observable data to separately identify the individual's tastes over outcomes and beliefs. This lack of identification complicates the interpretation of empirical findings, renders distinguishing between different theories difficult, and limits policy recommendations. For accurate predictions and effective policy, economists need models that are identifiable and refutable. In this paper, I offer such a model. My model implies conditions on observed choices that exhaustively characterise BDP. If choices abide by these conditions, then tastes and beliefs can be identified separately. If they do not, the theory

[1]For additional examples of information avoidance and distortion, see the surveys by Daniel & Hirshleifer (2015), Bénabou & Tirole (2016), and Golman et al. (2017).

is rejected. Such "if and only if" characterisation is absent in previous theories.

In theories of BDP, how an outcome is evaluated depends on the individual's beliefs. The lack of separation between these two concepts complicates their interpretation and identification because the way beliefs are formed may depend on tastes. Inferring tastes and beliefs from choices thus becomes a challenging task. The implication is that it is harder to obtain clear predictions. In addition to the lack of identification, current approaches make different assumptions about information processing, ranging from using Bayes rule (Köszegi, 2006) to assuming that individuals forget information (Bénabou, 2015) or even choose their beliefs at will (Brunnermeier & Parker, 2005). These assumptions are difficult to test. Therefore, the extant theories of BDP have two related limitations. They fail to make explicit the link between tastes and belief revision and tastes and beliefs are not uniquely identified.

In this paper, I propose a model of BDP that addresses these limitations. I introduce axioms on choices that characterise individuals as having tastes over their beliefs and belief revision rules related to their tastes. The theory identifies prior beliefs, tastes, and their related belief updating rules from observable choices.

I study the choices of contingent menus. Contingent menus are collections of acts available at a later time that are conditional on the realisation of an uncertain state of the world. Consider the following illustrative example. An investor can access her banking app to receive information about the market. If she accesses the app, she observes information such as prices or the available balance. She then chooses how much to invest or withdraw. Accessing the app or not represent two *contingent menus*, each associated to a probability distribution over menus of available acts. The realisation of the menu of acts that is finally available to the individual is conditional on an uncertain state of the world. Thus, observing an available menu of acts is informative for the investor. For example, if the investor opens the app and observes a high value of her portfolio, she receives information not only about the amount of money she can withdraw but also about the state of the market. If she chooses not to access the app, she receives no information about the market and her only available act is to do nothing.

The choice between contingent menus thus reflects intertemporal preferences over acts and information. As suggested by this illustrative example, in the general model of an individual's decision problem, there are three time periods. The individual first chooses one of the available *contingent menus*, each of which delivers a distribution of menus of acts conditional on the state of the world. Then, a menu of acts available to the individual realises, and the individual updates her beliefs on the basis of the observation of the realised menu. Finally, the individual chooses an act from the realised menu.

The reason why contingent menus allow for identification is as follows. Consider an individual who has to choose between Blackwell experiments. A Bayesian individual always prefers the most informative experiment. Instead, intuitively, individuals with different preferences over beliefs should prefer different experiments. When an individual with BDP chooses between experiments, she takes expectations over her belief-dependent expected utility, particularly over the posteriors that the experiments can induce. Under Bayesian updating, the mean of the posteriors induced by any experiment is equal to the prior. Therefore, if the individual maximises her belief-dependent expected utility under Bayesian updating, her choices over experiments cannot reveal preferences over distinct posteriors, as their expectations are the same. The intuition that different belief-dependent preferences correspond to different choices of experiments fails, as shown by (Eliaz & Spiegler, 2006; Liang, 2017). An immediate implication is that non-Bayesian updating overcomes this problem, as the mean of posteriors that are not the Bayesian update of a prior is not necessarily equal to the prior. However, non-Bayesian updating induces dynamic inconsistency. Choices over contingent menus allow me to identify how the individual manages such dynamic inconsistency, that is, how she deviates from Bayesian updating. In turn, non-Bayesian updating allows for the identification of preferences over beliefs.

Two crucial axioms characterise BDP preferences over contingent menus. First, I need an axiom similar to the independence of von Neumann & Morgenstern (2007), but weakened to obtain BDP. Here individuals are sensitive to the informational content of a realised menu of acts. Therefore, I formulate a weak version of independence that is satisfied only among contingent menus inducing the same inference about the uncertain state (Liang, 2017; Rommeswinkel et al., 2023). This weak version of independence induces dynamic inconsistency. After receiving information, the individual is tempted to deviate from her plan. This is because, since individual tastes depend on beliefs, the individual might distort them after receiving information. The second axiom states that the individual faces no temptation when the realised menu of acts comprises her preferred choice under both the Bayesian posterior and according to her preferred posterior. This axiom, which I call *strategic rationality for best likelihood*, ascribes the source of temptation to distorted belief updating due to BDP, and constitutes the main departure from the literature. The axiom is intuitive: it implies that the individual is not tempted when the information she receives is the one she prefers.

The main result is Theorem 1, which provides a functional representation of preferences over contingent menus. The choices that are consistent with the representation can be interpreted as follows. The individual chooses a contingent menu according to her preferences over acts in menu realisations and over the information it provides. She anticipates that once a menu of acts realises, she will distort her posterior beliefs away from the Bayesian update to satisfy her BDP.[2] She would like to choose according to the Bayesian posterior but is tempted to choose according to the distorted one. She solves the conflict by maximising a weighted average of two expected utilities, one under the Bayesian posterior and the other with the distorted one. The second result, Corollary 1, is an identification one. The choices over contingent menus allow me to uniquely pin down the individual's prior beliefs, tastes over outcomes and beliefs, distorted posterior beliefs, and the weight on the distorted utility.

Resorting to contingent menus as a primitive makes it possible to infer tastes over both beliefs and acts. This, in turn, overcomes the two issues that plague BDP highlighted above. First, it allows to test BDP by observing choices alone, in contrast to other extant theories. Second, BDP typically lead to dynamic inconsistency of behaviour in risky settings (Battigalli & Dufwenberg, 2022, p. 863). Dynamic inconsistency means that the choice the individual plans to take before receiving information is different from the one she wants to take after receiving information when she is tempted to act differently. Previous theories have relied on multiple-selves models. However, these models do not provide clear recommendations for welfare analysis, as a choice must be made regarding which of the selves' preferences are relevant. The current setting instead identifies the individual as the unit of choice, and her preferences can be subject to welfare analysis.

In sum, my model has two distinctive features compared with previous theories: it details the link between tastes over beliefs and belief updating, rather than assuming specific revision rules, and it relies only on observable choices to identify preferences, prior beliefs and revision rules, moving one step towards the empirical testing of BDP and non-Bayesian updating.

I compare the paper with the literature in the following paragraphs. Section 2 illustrates the primitives of the theory with two examples and provides instances of preferences satisfying the axioms. Section 3 details the general model and axioms. Section 4 presents the results, a functional representation of preferences satisfying the axioms and the uniqueness of the components of the functional form. Section 5 discusses the relationships with previous models. Section 7 concludes by discussing how the analysis sheds light on the interpretation of BDP.

[2]This interpretation relies on the sophistication of the individual at the ex ante stage. Cobb-Clark et al. (2022) provide evidence that a majority of time-inconsistent individuals exhibit at least partial sophistication.

Related literature. Unlike previous papers in decision theory, I allow tastes over beliefs to interact with tastes over outcomes and identify both prior beliefs and departures from Bayes rule. To the best of my knowledge, there have been three attempts to provide an axiomatic revealed preference foundation for BDP. Dillenberger & Raymond (2020) propose a model in which an individual has preferences over the probability of realisation of compound objective lotteries. In Liang (2017), the individual has preferences over the inference her choices induce in Anscombe & Aumann (1963)'s setting. Rommeswinkel et al. (2023) is similar to Liang (2017), except that the setting is that of Savage (1972). Given that it is an objective probability framework, Dillenberger & Raymond (2020) do not cover belief identification and updating. Moreover, despite working in a dynamic setting, they do not address inconsistency. Liang (2017) identifies beliefs when tastes over these are independent of tastes over outcomes. However, his setting is static and thus silent about belief revision. Rommeswinkel et al. (2023)'s model is instead dynamically consistent and identifies prior beliefs under the same separability assumption of Liang (2017), but belief updating is not addressed.

A novelty of the present paper is the study of contingent menus delivering different inferences about the state. Variation in the inference provided by contingent menus is the key to identifying preferences and belief revision rules. Therefore, the paper is related to the literature that studies menu choice to identify departures from subjective expected utility. Ozdenoren (2002) considers contingent menus of objective lotteries as primitives. Epstein (2006) and Epstein & Kopylov (2007) instead study contingent menus of Anscombe-Aumann acts. In all these papers, the state giving rise to the menu realisation is revealed; thus, preferences for information cannot be identified. The closest paper is Epstein (2006), who provides a model of non-Bayesian updating without considering BDP but does not study choice of information.

This paper is also related to the literature explaining empirical observations with BDP. Bénabou & Tirole (2016), Golman et al. (2017) and Battigalli & Dufwenberg (2022) are three surveys on motivated beliefs, information avoidance and psychological game theory, which are related to BDP. The distinction between the previous applied work and the present paper is testability. I follow the axiomatic approach and characterise behaviourally a class of preferences and revision rules. Previous papers have "if" results that can rationalise evidence but do not allow to distinguish between different theories. Moreover, I do not tailor the model to an application, and I do not commit to a specific psychological mechanism, taste or belief revision rule. Unlike psychological games, I consider a single individual and focus on identification rather than equilibrium in games. Finally, I address dynamic inconsistency via temptation and self-control, dispensing from previously employed multiple-selves or intrapersonal equilibrium approaches (Brunnermeier & Parker, 2005; Köszegi, 2010), which lead to problematic welfare analyses (Siniscalchi, 2011, p. 404) and are difficult to test, limiting policy recommendations.

2 Illustrative Examples

In this section, I develop two examples. The aim is both to illustrate the primitives of the model and to show how the theory explains some empirical observations.

Information distortion to reduce donations. A dictator game is an interaction between two individuals. One of them, the dictator, decides how much of a given amount to keep for herself and how much to transfer to the other individual, the recipient. It has been observed in the laboratory that dictators avoid receiving free information about how much of their transfer will arrive at the recipient to justify acting in a self-interested manner (Dana et al., 2007). Moreover, conditional on receiving favourable information, they transfer more (Grossman & van der Weele, 2013; Van der Weele, 2014). Similar instances of information avoidance are common.[3] I show how such attitudes towards information can be explained by non-Bayesian updating induced by BDP, and introduce a sketch of an experimental design to test the theory.

I consider a stochastic version of a dictator game. The dictator chooses how much to transfer to a recipient with whom she is coupled. The transfer is inefficient, and the receiver only receives a stochastic proportion of it. The dictator chooses whether to observe a signal on the efficiency of the transfer or not. If she observes the signal, she learns the likelihood $\ell(e)$ of the efficiency level having value e.

The dictator derives warm-glow from her expectation of the receiver outcome.[4] She would like to believe that the efficiency of the transfer is high to increase her warm-glow feelings. However, the higher the expected efficiency is, the higher the optimal transfer, which is a cost for the dictator. The dictator's tastes over transfers $x \in [0, 10)$ at likelihood ℓ are

$$u(x;\ell) = log(10 - x) + \sum_{e} p_{\ell}(e)\, ex, \tag{1}$$

where p_{ℓ} is the posterior belief over efficiency e after observing likelihood ℓ.[5] If the dictator

[3] See Section 3 in Golman et al. (2017) for multiple references.

[4] Niehaus (2014) proposed a similar preference in the context of charitable giving.

[5] The Bayesian posterior of the prior p after observing likelihood ℓ is $p_{\ell}(s) = \frac{\ell(s)p(s)}{\sum_{s'} \ell(s')p(s')}$.

chooses not to observe the signal, she chooses according to her prior beliefs, i.e., under the uninformative likelihood. The dictator maximises the sum of her material payoff, logarithmic in money, and her expectation of what the receiver receives. This expression deviates from expected utility because the information received in the form of the likelihood ℓ affects the taste over outcomes, not only the beliefs about their realisation. The optimal transfer at likelihood ℓ is

$$x^*(\ell) = 10 - \frac{1}{\sum_e p_\ell(e)\, e},$$

which is increasing in the expected efficiency. The function $u(x^*(\ell), \ell)$ is convex; therefore, its expectation over signal realisation is always greater than the prior. An individual computing expected utility with her Bayesian posterior and tastes in Equation (1) would always prefer to observe the signal and not avoid information.

Now consider a case in which the dictator distorts her beliefs after receiving the signal to maximise u. The likelihood ℓ^* of inducing the preferred beliefs satisfies the following

$$\ell^* \in \arg\max_{\ell} \max_{x} u(x; \ell).$$

In this case, the preferred beliefs give probability one to the highest level of efficiency and therefore induce a high transfer. If the dictator expects to distort her beliefs, she will avoid information, contrary to what she would do if she was Bayesian, to justify transferring less. Such extreme distortion is a particular case of the general model in the body of the paper.

The experimenter can allow the dictator to choose whether to observe the signal and to commit to a transfer conditional on the signal realisation before receiving it. If the dictator prefers to receive the signal and commit to a transfer conditional on it, it means that she wants to be informed but anticipates that she will distort the signal and act according to distorted beliefs if she has the chance.

Ostrich effect and excessive trading. This example examines how an investor's decision to seek information about the market is influenced by her tastes over the beliefs she holds. An investor chooses whether to check her financial portfolio. If she checks, she observes a signal about the state of the market and can invest or withdraw money. If she does not check, she receives no information and cannot invest or withdraw.

The investor enjoys believing that the market is in a good state and suffers when she does not. If she receives a bad signal, she suffers from negative news. Instead, if she receives a good signal, she overweights the evidence and develops overly optimistic beliefs. These distorted beliefs lead her to invest more than she would do on the basis of the Bayesian update of her prior beliefs. When choosing whether to check the portfolio, she weights the following factors: receiving bad news and suffering from it or receiving good news and acting on distorted beliefs.

If the investor has a low prior belief that the market is in a good state, she prefers not to check the portfolio to avoid unpleasant information. Instead, if she expects the market to be in good state, she may choose to check the portfolio to update her beliefs in a favourable direction rather than remaining uninformed. These behavioural patterns are well-documented (Daniel & Hirshleifer, 2015; Golman et al., 2017). The first pattern, known as the "ostrich effect" in finance, involves avoiding unpleasant information. The second pattern involves excessive motivation from overconfidence in belief formation, in this case leading to excessive trading.

I now introduce a utility function that consistent with the investor's choices and constitutes a particular case of the theory developed in this paper. The individual values both the monetary outcome of her decision and the information she receives, with these factors being separable. She values each unit of net monetary gain independently of her beliefs. She makes a choice at two time periods: first, she chooses whether to check the portfolio; and second, conditional on her previous choice and the signal she receives, she chooses how much to invest or withdraw.

The feasible distributions of net financial gains depend on the signal received. As an example, the investor can implement various investment strategies on the basis of the prices and available balance she observes. Therefore, a signal corresponds to a menu of feasible acts the individual can choose from, denoted by M. After observing the menu of feasible acts M as a signal, the individual updates her beliefs by combining her prior $p(s)$ with the likelihood $\ell_M(s)$ of state s being true given the observed menu M. The Bayesian posterior is $p_{\ell_M}(s)$. The individual then chooses an act f, an investment strategy, from the feasible set M, whose outcome f_s depends on the realisation of the state. Assume that there are three possible states: good (g), normal (n), and bad (b). The individual's tastes over monetary outcomes are $v(f_s)$. Her tastes over outcomes and information are represented by the following

$$u(f_s; \ell_M) = v(f_s) + \omega_g \ell_M(g) + \omega_n \ell_M(n),$$

where $\omega_g > \omega_n > 0$ are numbers representing how much the individual values observing greater likelihoods that the state is good and normal. The investor values a greater likelihood of the state being good than normal, which in turn is more valuable than observing a greater likelihood of the state being bad. A Bayesian individual with these tastes maximises the expectation of u computed via the Bayesian update

$$\sum_{s} p_{\ell_M}(s)\, u(f_s; \ell_M).$$

The most desirable information for the investor is a likelihood vector ℓ satisfying

$$\ell^* \in \arg\max_{\ell} \max_{x} \left[v(x) + \omega_g \ell(g) + \omega_n \ell(n) \right].$$

which is $\ell^*(g) = 1$ and $\ell^*(n) = \ell^*(n) = 0$.

If the investor checks her portfolio, she can receive one of two signals: a precise signal indicating conclusively that the market is in a bad state or an imprecise signal that rules out the bad state but does not allow her to distinguish between the good and normal states. The investor anticipates that upon receiving the imprecise signal, she will distort the likelihood to $\ell^*_M(g) = 1$, her preferred one according to her tastes u. Instead, the likelihood induced by the precise signal cannot be distorted, as there is strong conclusive evidence.

The optimal acts under the true and distorted likelihood are different because the second one is more optimistic and leads to greater investment. Theorem 1 from the general model establishes that, under some assumptions, the investor maximises a weighted sum of her expected utility under the true likelihood and under the distorted likelihood. Additionally, she incurs a cost depending on the utility difference between her choice and the optimal choice under the distorted posterior. The utility of choosing from the menu M at likelihood ℓ_M can be represented as follows:

$$\max_{f \in M} \left[\sum_{s} p_{\ell_M}(s)\, u(f_s; \ell_M) + \alpha_{\ell_M} \sum_{s} p_{\ell^*}(s)\, u(f_s; \ell^*_M) \right] - \max_{f' \in M} \alpha_{\ell_M} \sum_{s} p_{\ell^*_M}(s)\, u(f'_s; \ell^*_M), \tag{2}$$

where ℓ^*_M denotes the distorted vector of likelihoods after observing menu M as a signal, with $\ell^*_M(g) = 1$. The number $\alpha_{\ell_M} \geq 0$ represents the rate at which the cost of resisting temptation increases when the optimal choices under the Bayesian and distorted posteriors differ. The better an option is according to the distorted posterior, the higher the cost. It is uniquely identified in the model. When choosing whether to check her portfolio, the investor considers the probability of receiving each menu of feasible acts M as a signal and the corresponding indirect utility of choosing from it, given by the utility function above.

Consider the scenario where the investor can check her portfolio and commit to an investment strategy, for example by delegating to a financial advisor or an investment algorithm. In this case, she can receive information without the temptation to act on distorted beliefs. If the menu M observed as a signal is a singleton, the second and third terms in the indirect utility function in Equation (2) cancel out. As a result, preferences are represented by the expectation of tastes over outcomes and information computed with Bayes rule. Observing such preferences for commitment is a crucial test of the theory presented in the paper.[6]

These examples suggest how the model components are identified. Choosing from singleton menus does not give rise to temptation to distort beliefs. Therefore, choosing from singleton menus is a standard choice problem, and the results of Anscombe & Aumann (1963) helps to identify tastes over outcomes and beliefs after each signal. Observing choices over information and menus of acts feasible at a later time, transfers in the first example and investment strategies in the second, allow identifying a demand for commitment due to distortion of beliefs in a specific direction. Therefore, the data needed to test the model are choices over information and sets of available acts conditional on the information received. I refer to such an object of choice as a contingent menu, introduced in the next section.

3 Model

An individual faces a dynamic decision problem in which she makes two choices. First, she chooses an act that maps states to menus of acts from which she can later select. Then, she chooses an act from the realised menu, which maps states to outcomes. Upon observing the menu from which she can choose, the individual infers information about the state. This information is relevant for her choice of an act from the realised menu.

The decision problem involves three time periods $t = 0, 1, 2$. There is a finite set of uncertain states $\mathcal{S}$. At time 0, the individual chooses a contingent menu, a mapping from uncertain states to finite probability distributions over menus of acts she will choose from at a later time. A generic menu realisation is $M \in \mathcal{M}$. A contingent menu is $F : \mathcal{S} \to \Delta^\circ(\mathcal{M})$, where $\Delta^\circ(\mathcal{M})$ is the set of probability distributions over $\mathcal{M}$ with finite support. At time 1, a menu M realises. At time 2, the individual chooses an act from the menu M. An act is a mapping between states and outcomes $f : \mathcal{S} \to \Delta(X)$, where X is a compact metric set, and $\Delta(X)$ is

[6]Relatedly, Derksen et al. (2024) show that medical appointments are effective commitment devices that significantly increase the probability of individuals at health risk getting tested.

the set of probability distributions over X.[7] A menu M is therefore a closed nonempty set of acts.[8] The outcome induced by act f when state s realises is $f_s \in \Delta(X)$.

Observing the menu realisation M from the contingent menu F at time 1 is informative for the individual. In fact, a contingent menu is a Blackwell experiment with menus as signals. Denote the probability that menu M is realised from contingent menu F in state s with $F_s(M)$. To capture the informational content of a contingent menu F, I define the normalised likelihood (henceforth likelihood) of state s after realisation of menu M

$$\ell_{M,F}(s) := \frac{F_s(M)}{\sum_{s'} F_{s'}(M)}. \tag{3}$$

After having observed likelihood ℓ, the individual knows that the state is in the event

$$S_\ell := \{s \in \mathcal{S} \mid \ell(s) > 0\}.$$

Throughout, I assume that only choices of contingent menus are observable. The main result of this paper identifies all the components of the following model relying on these choices. In particular, I introduce the likelihood because I assume that the beliefs of the individual are not observable and will therefore be inferred from her choices. I denote preferences over contingent menus with $\succsim$. Theorem 1 in Section 4 provides the conditions on choices over contingent menus that yield the following representation of preferences.

Representation. The main result of the paper, Theorem 1, shows that the axioms in Section 3.1 are equivalent to the following representation. The individual has belief-dependent tastes over outcomes and likelihoods represented by a utility function $u(x;\ell)$, linear in mixtures of outcomes $x \in \Delta(X)$, jointly continuous, bounded, and nonconstant for each ℓ. The dependency of u on the likelihood ℓ is the main departure from expected utility. The individual evaluates outcomes differently depending on the information she receives.

The individual acts as if she has a full support prior over states p. The Bayesian posterior of the prior p after observing the likelihood ℓ is p_ℓ, where for each state s

$$p_\ell(s) := \frac{\ell(s)\,p(s)}{\sum_{s'} \ell(s')\,p(s')}.$$

The time 2 expected utility of act f computed with the Bayesian posterior at likelihood ℓ is

[7]The set of lotteries $\Delta(X)$ is compact metric under the weak convergence topology.

[8]The set of menus $\mathcal{M}$ is compact metric under the Hausdorff metric (Aliprantis & Border, 2006, Theorem 3.85).

$$\sum_{s} p_{\ell}(s)\, u(f_s;\ell)\,. \tag{4}$$

The individual is tempted to act according to a distorted posterior. This posterior is the one obtained under the preferred likelihood consistent with the event S_ℓ. For any event $S \subseteq \mathcal{S}$, define[9]

$$\ell_S^* \in \arg\max_{\ell\in\Delta(S)} \max_{x\in\Delta(X)} u(x;\ell)\,. \tag{5}$$

The distorted likelihood ℓ_S^* only assigns positive probability to states in event S. Once a state has probability 0, its probability cannot be distorted. Moreover, the distortion only depends on the event, not on the true likelihood. There is no guarantee that there is a unique likelihood satisfying Equation (5). However, the model allows the identification of one likelihood among those satisfying Equation (5) from choices over contingent menus.

The time 2 expected utility of act f computed with the distorted posterior at event S is

$$\sum_{s} p_{\ell_S^*}(s)\, u(f_s;\ell_S^*)\,.$$

The distorted likelihood influences expected utility through two channels. It affects tastes over outcomes $u(x;\ell_S^*)$ and posterior beliefs, which are the Bayesian update of the prior p under the distorted likelihood ℓ_S^*.

At time 2, after having observed the menu realisation M from contingent menu F, the individual chooses an act f from M to maximise a weighted combination of the expected utilities under the true and distorted likelihood. Moreover, she suffers a cost proportional to the utility difference between the chosen act and the optimal act under the distorted likelihood. For each event S, the utility representation over menus at each likelihood ℓ is as follows:

$$\begin{aligned}\mathcal{U}(M;\ell) = &\max_{f\in M}\left[\sum_{s} p_{\ell}(s)\, u(f_s;\ell) + \alpha_\ell \sum_{s} p_{\ell_{S_\ell}^*}(s)\, u\left(f_s;\ell_{S_\ell}^*\right)\right] \\ &- \alpha_\ell \max_{f'\in M} \sum_{s} p_{\ell_{S_\ell}^*}(s)\, u\left(f'_s;\ell_{S_\ell}^*\right),\end{aligned} \tag{6}$$

where the positive number $\alpha_\ell \geq 0$ represents the weight assigned to the distorted expected utility. When choosing act f from menu M after realisation of the likelihood ℓ, the utility cost of temptation over menus is the difference between the second and the third terms in Equation

[9]Such likelihoods always exist since u is continuous and both $\Delta(X)$ and $\Delta(S)$ for each event S are compact.

(6). Specifically, it is the utility difference between the chosen act and the optimal act under the distorted likelihood. The representation implies that, for each event S, when the true likelihood coincides with ℓ^*_S, the preferred likelihood, there is no temptation. When the individual receives the information she prefers, there is no reason to distort it.

I can now describe preferences over contingent menus. The individual chooses the contingent menu F anticipating the indirect utility $\mathcal{U}(M; \ell_{M,F})$ from each possible menu realisation M, so that each contingent menu F is evaluated by the expected utility

$$\mathscr{U}(F) = \sum_M \sum_s p(s) F_s(M) \mathcal{U}(M; \ell_{M,F}). \tag{7}$$

To summarise, the model is as follows. When choosing the contingent menu F, the individual anticipates that her BDP will lead her to update prior beliefs p deviating from Bayes rule after observing the menu realisation M. This deviation is represented by the distortion of the true likelihood ℓ to ℓ^*, which leads to updating the prior beliefs with Bayes rule via ℓ^*. The interpretation is that, ex ante, she would like to choose from any menu to maximise her expected utility under the Bayesian update. However, she is tempted to maximise her expected utility under the Bayesian update of the prior given the distorted likelihood. She is sophisticated and foresees the temptation of choosing according to the distorted posterior, influencing both preferences and beliefs. Thus, there is a trade-off between acting according to the Bayesian and distorted posterior beliefs.

3.1 Axioms

In this section, I list the axioms on the preference relation $\succsim$ over contingent menus yielding the representation in Section 3. I begin with the standard axioms that allow a continuous utility representation of preferences over contingent menus to be obtained.

Axiom 1. ***(Order)****. The ranking $\succsim$ is complete and transitive.*

Axiom 2. ***(Continuity)****. For each contingent menu F the sets*

$$\{F' \mid F' \succsim F\} \text{ and } \{F' \mid F' \precsim F\}$$

are closed.

I now introduce an axiom structuring the attitude to information. Since in this model the individual has preferences for the information she receives, contingent menus with the same informativeness play a special role. First, I define the support of a contingent menu F

$$\mathcal{M}_F := \{M \in \mathcal{M} \mid F_s(M) > 0 \text{ for some } s \in \mathcal{S}\}.$$

Recall that a likelihood is a probability distribution over states defined in Equation (3)

$$\ell_{M,F}(s) := \frac{F_s(M)}{\sum_{s'} F_{s'}(M)}. \tag{3}$$

DEFINITION 1. ***(Identical Inference (II))*** *Two contingent menus F and F' satisfy* ***identical inference*** *if, for each menu $M \in \mathcal{M}_F \cap \mathcal{M}_{F'}$, their likelihood is the same $\ell_{M,F} = \ell_{M,F'}$.*

Two contingent menus F, F' satisfying II have the property that, when a menu M is realised from a probabilistic mixture of them, inference about the state is the same regardless of whether it comes from F or F'. To state independence, I first define the relevant mixture operations. As usual, a mixture of two acts delivers in each state a probability distribution that is the mixture of the one induced by the two acts. For any two acts f, f', state s and outcome x

$$(\lambda f + (1-\lambda) f')_s(x) = \lambda f_s(x) + (1-\lambda) f'_s(x).$$

As is standard in the menu choice literature, a mixture of two menus is a menu of mixed acts, one of which is in the first menu and the other in the second menu. For any two menus M and M' and $0 \leq \lambda \leq 1$,

$$\lambda M + (1-\lambda) M' = \{\lambda f + (1-\lambda) f' \mid f \in M, f' \in M'\}.$$

I now define mixtures of contingent menus. The contingent menu $\lambda F + (1-\lambda) F'$ delivers a distribution of menus conditional on each state s, which is a mixture of F_s and F'_s. For any two contingent menus F, F', state s and menu M

$$(\lambda F + (1-\lambda) F')_s(M) = \lambda F_s(M) + (1-\lambda) F'_s(M).$$

I now impose a weak version of independence that holds only among II contingent menus.

AXIOM 3. ***(II Independence).*** *For all $0 < \lambda \leq 1$ and contingent menus F, F', F'' such that F and F'' satisfy II and F' and F'' satisfy II, $F \succsim F'$ if and only if $\lambda F + (1-\lambda) F'' \succsim \lambda F' + (1-\lambda) F''$.*

This axiom constrains preferences to depend on the realised likelihood. The intuition for the axioms is as follows. Under expected utility, independence holds to induce preferences that are linear in probabilities. However, if preferences depend on the information received, the standard independence axiom is not appropriate. This is because mixing two contingent menus affects the likelihoods that they induce for each of their menu realisation. Preferences over information should not be linear in such mixtures of likelihoods. II Independence imposes indifference only for mixtures of contingent menus inducing the same likelihood from menus in their common support.

Restricted to II contingent menus, the intuition for independence is the standard one in the menu choice literature. Adapted to the present setting, it is as follows. Consider a lottery over contingent menus delivering F with probability λ and F' with probability $1-\lambda$. The intuition for independence suggests that $F \sim F'$ iff such a mixture is indifferent to both F and F'. Once justification for indifference between probabilistic mixtures and $\lambda F_s + (1-\lambda) F'_s$ is provided, the intuition for independence is complete. Under the latter, first, the state is realised, and the individual updates her beliefs and then chooses from the available menu of mixed acts. Under probabilistic mixtures, randomisation among contingent menus is performed before the individual's choice. Hence, indifference between the two amounts to indifference to the timing of resolution of these sources of uncertainty. I briefly comment on such indifference in the conclusion.

Identify with y the contingent menu delivering in every state the singleton menu containing the act yielding the outcome y with probability 1. To avoid trivial cases, I assume the following.

AXIOM 4. *(**Nondegeneracy**). There exist outcomes y, y' in X for which $y \succ y'$.*

Next, I adapt to the present setting the Set-Betweenness axiom of the menu choice literature. The intuition behind the axiom is that the individual prefers not to expand the available menu with ex-ante suboptimal acts, as these create temptation. A new notation is needed. For any contingent menu F, menu M in its support and menu M' outside its support, I denote with $F_{M\to M'}$ a contingent menu equivalent to F except that any realisation of M is substituted with M'. The menu M' should not be in the support of F; otherwise, $\ell_{M',F_{M\to M'}}$ would not be identical to $\ell_{M,F}$.

AXIOM 5. *(**Set-Betweenness**). For all contingent menus F and menus M, M'*

$$F \succsim F_{M\to M'} \Rightarrow F \succsim F_{M\to M\cup M'} \succsim F_{M\to M'}.$$

The rationale for this assumption is the same as in the menu choice literature (Gul & Pesendorfer, 2001), except it holds conditional on observing a menu realisation. The preference $F \succsim F_{M \to M'}$ indicates that the individual would rather choose from M than from M', all else equal. Temptation cannot increase utility; hence, the individual prefers not to expand M with M', which contains ex-ante dominated options. Since the two contingent menus are otherwise equivalent at the ex ante stage, the ranking in the axiom follows.

The next axiom is the main departure from the literature. Its role is to ascribe the source of temptation to belief distortion due to BDP. The idea is that the individual should not distort her beliefs when observing a likelihood that satisfies her BDP preferences. To state the axiom, I must define such a best likelihood. For this purpose, a few definitions are needed.

The following is the set of all contingent menus that induce likelihood ℓ whenever menu M realises

$$\mathcal{C}_{\ell}^{M} := \{F \mid \ell_{M,F} = \ell\}.$$

Then, I define the set of preferred outcomes at likelihood ℓ[10]

$$X_{\ell} := \left\{ x \in \Delta(X) \,\middle|\, F \succsim F_{\{x\} \to \{x'\}} \text{ for all } x' \in \Delta(X) \text{ and some } F \in \mathcal{C}_{\ell}^{\{x\}} \right\}.$$

Owing to II Independence, the ranking between any two menus M and M' does not depend on the specific contingent menu F, as long as $\ell_{M,F} = \ell$. Therefore, II Independence implies that "for some" in the definition of X_{ℓ} is equivalent to "for all". A generic element of X_{ℓ} is x_{ℓ}. For illustration, I will show that in terms of the representation in Section 3 each x_{ℓ} satisfies the following:

$$x_{\ell} \in \arg\max_{x \in \Delta(x)} u(x; \ell).$$

Fix a collection of outcomes $(x_s)_{s \in \mathcal{S}}$. For each ℓ, construct the contingent menu F^{ℓ} such that $F_s^{\ell}(\{x_{\ell}\}) = \ell(s)$ and $F_s^{\ell}(\{x_s\}) = 1 - F_s^{\ell}(\{x_{\ell}\})$ for all s.[11] For each S, define the likelihoods

$$\ell_S^{*} \in \left\{ \ell \in \Delta(S) \,\middle|\, F^{\ell} \succsim F^{\ell'} \text{ for all } \ell' \in \Delta(S) \right\}. \tag{8}$$

[10]Since $\succsim$ is continuous and $\Delta(X)$ is compact, each X_{ℓ} is nonempty.

[11]An example of the construction of such a contingent menu is shown in Appendix B.

I show in Theorem 1 that any ℓ^*_S satisfies Equation (5) for each S:

$$\ell^*_S \in \arg\max_{\ell \in \Delta(S)} \max_{x \in \Delta(X)} u(x;\ell). \tag{5}$$

These likelihoods can be interpreted as follows. Say the individual can choose an outcome in $\Delta(X)$, whose realisation does not depend on the state. For each event S, the likelihood ℓ^*_S is the likelihood that the individual would prefer to observe at that event. The likelihood determined via this procedure reflects the individual's preferences over information when it is not instrumental for choice.

One more piece of notation is necessary to state the axiom. Define for each menu M and likelihood ℓ the set of acts

$$\mathcal{F}_{M,\ell} := \left\{ f \in M \,\middle|\, F \succsim F_{\{f\}\to\{f'\}} \text{ for all } f' \in M \text{ and some } F \in \mathcal{C}^{\{f\}}_{\ell} \right\}.$$

Fix a menu M and a likelihood ℓ induced by it. Then, $\mathcal{F}_{M,\ell}$ is interpreted as the set of ex ante best acts in M at that likelihood. If the individual could commit, she would always choose acts from this set.[12] To illustrate, I will show that, in terms of the representation, acts in $\mathcal{F}_{M,\ell}$ are those maximising Equation (4) and therefore satisfying for each ℓ and M the following

$$f \in \arg\max_{f' \in M} \sum_s p_\ell(s)\, u(f'_s;\ell).$$

I now state the axiom.

AXIOM 6. ***(Strategic Rationality for Best Likelihood (SRBL))****. For each:*

- *couple of menus M, M';*
- *contingent menu F such that $\ell_{M,F} = \ell$;*

if $\mathcal{F}_{M\cup M',\ell} \cap \mathcal{F}_{M\cup M',\ell^*_{S_\ell}} \neq \emptyset$ for at least one $\ell^*_{S_\ell}$, then

$$F \succsim F_{M\to M'} \Rightarrow F \sim F_{M\to M\cup M'}.$$

[12]Each menu M is a subset of the set of acts $\Delta(X)^{\mathcal{S}}$. Since $\mathcal{S}$ is finite, $\Delta(X)^{\mathcal{S}}$ is the cartesian product of compact spaces. By Theorem 2.61 in Aliprantis & Border (2006, p. 52), the cartesian product of compact spaces is compact. A menu M is thus a closed subset of a compact space, and is therefore compact. By compactness of M and continuity of preferences $\succsim$, each $\mathcal{F}_{M,\ell}$ is nonempty.

The intuition for the axiom is as follows. First, notice that the axiom implies that at the preferred likelihoods ℓ^*_S there is never temptation, as the antecedent is always satisfied. There is no reason to distort beliefs if They are the desired ones. However, the axiom imposes more. There is no temptation whenever the optimal choice under the true and preferred likelihoods coincides. The intuition is that, in this case, there is no trade-off between acting according to the true or distorted likelihood, and therefore no demand for commitment.

It is instructive to consider what would happen if II Independence is strengthened to hold among all contingent menus. In fact, the interaction between II Independence and SRBL allows interpretation of the distorted belief updating in the functional form as coming from BDP and not from other cognitive phenomena. Assume that the individual observes likelihood ℓ, which is different from one of her preferred likelihoods $\ell^*_{S_\ell}$ at that event. Say the same act in the menu M at her disposal is optimal under both likelihoods. Then, she faces no temptation and picks this act. When II Independence is strengthened to hold for all contingent menus, regardless of their informational content, the individual has no BDP. In terms of the representation, this means that u only depends on outcomes, not on likelihoods. Preferences over likelihoods are flat, and SRBL implies that there is never temptation. The classical version of independence and SRBL together imply that the individual is always strategically rational and the model collapses to expected utility with Bayesian updating. SRBL traces the source of temptation to having an optimal choice that is different under the true and preferred information, and its antecedent holds for all likelihoods when there are no BDP.

The next axiom is an adaptation of state independence to the current setting. The notation fsf' indicates an act equivalent to f in state s and to f' in all states $s' \neq s$. For each state s and menus M, M', define the menu $MsM' := \{fsf' \mid f \in M, f' \in M'\}$. The menus of outcomes are denoted with $L \subseteq \Delta(X)$.

AXIOM 7. ***(State Independence)***. *For all contingent menus F, menus L, L', M and states s, s'*

$$F \succsim F_{LsM \to L'sM} \;\Rightarrow\; F \succsim F_{Ls'M \to L's'M}.$$

The contingent menus F and $F_{LsM \to L'sM}$ are equivalent except in one realisation. The first offers a choice of outcomes from L in state s, whereas the second from L' in the same state. The axiom requires that the ranking of the two contingent menus is preserved when the state is changed. The intuition is that the individual's preferences over menus of outcomes are independent of the state in which the lottery realises.

Finally, I assume full support.

AXIOM 8. *(**Full Support**). For each state s, there exist contingent menus F and F' such that for all menus M it holds that $F_{s'}(M) = F'_{s'}(M)$ for each $s' \neq s$ and $F \nsim F'$.*

If two contingent menus are always indifferent whenever they are equivalent in each state except one, then that state can be omitted without loss of generality.

4 RESULTS

Representation. The main result of this paper links axioms to the utility representation in Equations (6) and (7). The proofs are in Appendix A. I report the utility representation and its properties.

DEFINITION 2. *A ranking $\succsim$ over contingent menus is a BDP if it is represented by Equations* (6) *and* (7)

$$\mathscr{U}(F) = \sum_M \sum_s p(s) F_s(M) \mathcal{U}(M; \ell_{M,F}), \tag{7}$$

$$\begin{aligned} \mathcal{U}(M;\ell) = &\max_{f \in M} \left[\sum_s p_\ell(s) u(f_s;\ell) + \alpha_\ell \sum_s p_{\ell^*_{S_\ell}}(s) u\left(f_s;\ell^*_{S_\ell}\right) \right] \\ &- \alpha_\ell \max_{f' \in M} \sum_s p_{\ell^*_{S_\ell}}(s) u\left(f'_s;\ell^*_{S_\ell}\right), \end{aligned} \tag{6}$$

where:

1. *$u : \Delta(X) \times \Delta(\mathcal{S}) \to \mathbb{R}$ is linear in mixtures in $\Delta(X)$, jointly continuous, bounded and nonconstant for each ℓ;*
2. *p is a full-support probability distribution over $\mathcal{S}$;*
3. *$\alpha_\ell \geq 0$ for each likelihood ℓ;*
4. *p_ℓ is the Bayesian posterior of p under likelihood ℓ;*
5. *ℓ^*_S satisfies Equation* (5) *for each S*

$$\ell^*_S \in \arg\max_{\ell \in \Delta(S)} \max_{x \in \Delta(X)} u(x;\ell). \tag{5}$$

I can now state the main result.

THEOREM 1. *The ranking $\succsim$ over contingent menus satisfies the axioms Order, Continuity, II Independence, Nondegeneracy, Set-Betweenness, SRBL, State Independence* and *Full Support if and only if it is a BDP.*

If an individual's preferences over contingent menus satisfy the axioms in the statement of Theorem 1, she behaves as if she anticipates distorting her beliefs to satisfy her BDP and act according to such distorted beliefs.

Uniqueness. I describe the uniqueness properties of the representation. Denote with $\ell^* = (\ell^*_S)_{S\in 2^S}$ a collection of preferred likelihoods satisfying condition (8), one for each event. Denote with p_ℓ the vector of posterior beliefs induced by observing likelihood ℓ, one for each s. Finally, $\alpha = (\alpha_\ell)_{\ell\in\Delta(S)}$.

COROLLARY 1. *Let (u, p, α, ℓ^*) represents $\succsim$, then $(u', p', \alpha', \ell'^*)$ also represents $\succsim$ if and only if there exists $(a, b) \in \mathbb{R}_{++} \times \mathbb{R}$ such that*

$$u' = au + b \quad and \quad p' = p.$$

For each likelihood ℓ, if $\alpha'_\ell \neq \alpha_\ell$, then $\ell = \ell^*_{S_\ell}$. Last, $\ell'^*_S = \ell^*_S$ for each event S.

It is instructive to compare the uniqueness properties of the representation with previous results from the literature on BDP. Eliaz & Spiegler (2006) and Liang (2017) show that preferences over posterior beliefs unique up to monotonic transformations cannot be identified from choices of information and acts alone. The lack of uniqueness is because the mean of the posteriors is equal to the prior beliefs. In the language of the current model, they establish that the function $\mathcal{U}'$ represents the same ranking as $\mathcal{U}$ if and only if there exists $(a, b, c) \in \mathbb{R}_{++} \times \mathbb{R} \times \mathbb{R}^S$ such that for each likelihood ℓ

$$\mathcal{U}'(\cdot; \ell) = a\mathcal{U}(\cdot; \ell) + b - \sum_s c(s)\, p_\ell(s)\,. \tag{9}$$

When taking expectations over $\mathcal{U}$, the term $\sum_s c(s)\, p_\ell(s)$ averages to a constant for all likelihoods.[13] Eliaz & Spiegler (2006) provide examples showing how the lack of uniqueness prohibits the identification of preferred beliefs. An individual's choices of information can reveal only preferences for probability distributions which mean is the prior. Non-Bayesian updating is thus responsible for the identification result in Corollary 1. The uniqueness properties of $\mathcal{U}$

[13]The algebra is as follows:

are inherited by the function u. Owing to the functional form of $\mathcal{U}$, the term $\sum_s c(s)\, p_\ell(s)$ must necessarily be null for Equation (9) to hold. Since $u(\cdot; \ell^*_S)$ appears but the average of $\sum_s c(s)\, p_{\ell^*_S}(s)$, with the distorted likelihood, is not a constant, c must be 0 for the transformation to represent the same ranking.

Identification of the model components allows elaborating on the behavioural meaning of α_ℓ. First, define conditional preferences at likelihood ℓ as follows:

$$M \succsim_\ell M' \text{ if } F \succsim F_{M\to M'} \text{ for some } F \text{ such that } \ell_{M,F} = \ell,$$

where "for some F" is equivalent to "for all F" under the axioms. The ranking $\succsim_\ell$ is represented by $\mathcal{U}(\cdot;\ell)$. Consider act f and outcomes $x, x' \in \Delta(X)$ such that for some ℓ

$$\{f\} \succ_\ell \{f, x\} \succ_\ell \{x\}\ ,\ \{f\} \succ_\ell \{f, x'\} \succ_\ell \{x'\} \text{ and } \{x\} \nsim_\ell \{x'\}\,.$$

In other words, outcomes x and x' are tempting when choosing from menus $\{f, x\}$, $\{f, x'\}$ and x and x' are not indifferent at likelihood ℓ. Then

$$\mathcal{U}(\{f\};\ell) - \mathcal{U}(\{f,x\};\ell) = \alpha_\ell\left(u(x;\ell) - \sum_s p_\ell(s)\, u(f_s;\ell)\right),$$
$$\mathcal{U}(\{f\};\ell) - \mathcal{U}(\{f,x'\};\ell) = \alpha_\ell\left(u(x';\ell) - \sum_s p_\ell(s)\, u(f_s;\ell)\right).$$

Subtracting the two equations, the following expression for α_ℓ yields

$$\alpha_\ell = \frac{\mathcal{U}(\{f,x\};\ell) - \mathcal{U}(\{f,x'\};\ell)}{u(x;\ell) - u(x';\ell)}\,.$$

Theorem 9 in Gul & Pesendorfer (2001) allows interpreting α_ℓ as a measure of self-control. The ranking $\succsim_\ell$ exhibits less self-control than $\succsim'_\ell$ if for all menus of acts M and M', the ranking

---

$$\begin{aligned}\sum_M \sum_s p(s)\, F_s(M) \sum_{s'} c(s')\, p_{\ell_{M,F}}(s') &= \sum_M \sum_s p(s)\, F_s(M) \sum_{s'} c(s')\, \frac{F_{s'}(M)\, p(s')}{\sum_{s''} F_{s''}(M)\, p(s'')} \\ &= \sum_M \sum_s c(s)\, F_s(M)\, p(s) \\ &= c\,.\end{aligned}$$

$M \succ_\ell M \cup M' \succ_\ell M'$ implies that the same must hold for $\succsim'_\ell$. In the current setting, self-control is relative to acting according to distorted beliefs. Therefore, the interpretation of α_ℓ can be adapted as a measure of the strength of motivated reasoning. As in Epstein (2006), the number α_ℓ is an absolute measure. The difference $\mathcal{U}(\{f\};\ell) - \mathcal{U}(\{f,x\};\ell)$ is the utility cost of self-control when the lottery x is available but the act f is chosen. Then, α_ℓ, is the rate at which the cost of resisting temptation increases as x improves, as measured by $u(x;\ell)$. It is the marginal cost of self-control at likelihood ℓ.

5 DISCUSSION

In this section, I discuss the model and its relationship with the previous literature. First, it is instructive to compare this model to Epstein (2006) and Gul & Pesendorfer (2001), which are the closest in the decision theory literature. In the present model, the distortion of the likelihood induces a change in both beliefs, via non-Bayesian updating, and tastes, trough BDP. A change in both tastes and beliefs constitutes a departure from the previous literature. In Epstein (2006), temptation arises because of non-Bayesian updating due to cognitive biases, not BDP. The individual does not change tastes as represented by u but suffers from updating biases. She is thus tempted to act according to her biased posterior beliefs. In Gul & Pesendorfer (2001), instead, temptation arises because of a change in tastes. After observing her menu, the individual is tempted to choose according to new tastes v. In the present model, both sources of temptation are present and linked to each other. The individual distorts her posterior beliefs *because* her tastes u depend on them. These distortions have structure, as distorted tastes are those the individual would have if the true likelihood was her preferred one ℓ^*.

Second, I discuss the relationships with other models of BDP and non-Bayesian updating. COmpared with previous models, there are two main distinctions. First, I do not take a stance on the cognitive process underlying belief distortion. The main advantage of this methodological stance is that the object of choice is observable and that predictions do not rely on a psychological interpretation. Instead, previous studies have resorted to various cognitive assumptions. As an example, in Brunnermeier & Parker (2005) and Köszegi (2006), the individual chooses her beliefs. Assuming that beliefs are chosen makes theories hard to test and obfuscates their revealed preference foundations. Spiegler (2008) shows that in the two models above the individual violates Independence of Irrelevant alternatives. The implication is that the individual's ranking of options depends on the set of options itself, and predictions vary significantly. A second relevant model is that of Bénabou (2015) and Bénabou & Tirole (2016), where the individual chooses the probability with which she forgets a signal and updates her beliefs by recalling such probability. Their model features multiple-selves with different preferences playing a game whose equilibrium is a forgetting strategy. The drawback of such a modelling approach is that the relevant unit of choice is not a single individual, which complicates the interpretation of the theory and leaves degrees of freedom to conduct welfare analyses. Second, in the present model non-Bayesian updating is not disjointed from BDP. The models above instead make two disjointed assumptions: that individuals' well-being depends on their beliefs and how they update their beliefs. Instead, here knowledge of the function u implies knowledge of how beliefs are distorted.

The informational parsimony in assumptions does not come at a cost to consistency with the empirical evidence. The robust stylised fact that individuals update suitably when facing good news but fail to properly account for bad news is consistent with the model.[14] When observing a likelihood different from the preferred likelihood, individuals exhibit non-Bayesian updating, but they do not when they face what they want to hear.

I consider the posterior obtained by updating conditional on the likelihood ℓ^*, distorted compared with the Bayesian update of the subjective prior p. This is in contrast to previous literature, in which individuals distorted their beliefs compared with an objective probability distribution that is considered "true" or has an empirical counterpart observable by the modeller (Brunnermeier & Parker, 2005; Yariv, 2002). The model thus reconciles motivated belief updating to subjective Bayesianism in the tradition of Savage (1972). Such distortion operates by interpreting the objective likelihood ℓ as ℓ^*, which is arguably a mistake. I provide a second interpretation of the model in the concluding remarks relying on a more "rational" view of such a decision criterion.

The particular choice of preferred likelihood requires justification. First, why consider the preferred likelihood given that the choice is from the menus of objective lotteries? An alternative is to identify the preferred likelihood when the choice is from among all the acts. The outcome of acts depends on the realisation of the uncertain state. Therefore, such a preferred likelihood would reflect not only preferences over beliefs but also an evaluation of the instrumental value of beliefs. To illustrate why this procedure is conceptually confusing, consider an individual with no BDP. Her "preferred beliefs" would be those which put all the weight

[14] See Eil & Rao (2011), Garrett & Sharot (2017), Möbius et al. (2022), Drobner & Goerg (2024), among others.

on states inducing her preferred outcome under some act. The reason for limiting choice to objective lotteries is thus to identify only the BDP component of the preferred likelihood, not the instrumental one. To see this intuition formally, consider that if I were to strengthen II Independence to hold for all contingent menus, then for any event S, all likelihoods are indifferent, as discussed above. If the individual has no BDP, her choices from objective lotteries do not depend on information and preferences on likelihoods are flat. Assume that I considered preferred likelihoods given that the choice is from some menu of acts M. The reasoning above would not hold; the preferred likelihoods of an individual satisfying independence are those that induce her preferred outcomes with probability 1 for some act. Therefore, SRBL together with Independence would imply that an individual with no BDP is tempted to act according to such degenerate beliefs, even if there is no reason to have them. Independence and SRBL would not deliver standard expected utility with Bayesian updating.

Second, why consider the preferred likelihood when the choice is from all objective lotteries and not one conditional on each possible menu of objective lotteries? This procedure would make the preferred likelihood depend on the event and on the outcomes in X that could be induced by the available acts and thus on the menu realisation. The model would be more complex without a clear gain in scope. Moreover, the representation would be weaker, as its components depend directly on the contingent menu. Regardless, the expression of SRBL does not depend on the definition of best likelihood and can accommodate other interpretations.

6 Application: Polarisation

In this section, I develop a simple application of the model to show how BDP can lead to belief polarisation. I assume a sender wants to persuade a receiver with BDP to take a specific action. The sender offers a menu of Blackwell experiments and individual with the same tastes over outcomes, but different tastes over beliefs, choose different experiments. The example is a simple variant of the judge and prosecutor example in Kamenica & Gentzkow (2011) and has various interpretations.

There is a binary state space $\mathcal{S} = \{0, 1\}$. An individual chooses an action $a \in \{0, 1\}$. The individual wants to match the state. She also has an identity $i \in [0, 1]$, representing a belief she would like to hold. The sender and the individual have a common prior of $p(0) = 7/10$. The individual's utility of choosing action a at state s and likelihood ℓ is

$$w(a;s,\ell) = -(a-s)^2 - (p_\ell - i)^2,$$

where p_ℓ is the posterior belief of state 1 at likelihood ℓ.[15]

Assume the sender wants to steer the individual toward choosing $a = 1$. The sender can choose experiments, mappings between states and distribution over action recommendations $E : \mathcal{S} \to \Delta(\{0,1\})$ and commits to reporting the signal, as in Bayesian Persuasion. When $i = 0$, the individual has the same preferences as in Kamenica & Gentzkow (2011), and the optimal experiment is the following:

	$E_s(a)$	0	1
$7/10$	0	$4/7$	$3/7$
$3/10$	1	0	1

Table 1: Optimal experiment for $i = 0$.

When $i > 1/2$, the individual prefers to have belief closer to 0 than to 0. In this case, the optimal experiment is the following:

	$E_s^2(a)$	0	1
$7/10$	0	p	$1-p$
$3/10$	1	0	1

Table 2: Optimal experiment for $i \in (1/2, 1]$.

where

$$p = \frac{4 + 10\alpha(1-2i)}{7} \quad \text{and} \quad p \neq 1.$$

The sender can induce the individual to choose action $a = 1$ more often when the individual has $i > 1/2$ compared to what she can do in the standard case when $i = 0$. Moreover, the sender can do better when α, the strength of motivated reasoning, is higher.

[15] In the language of the model in the paper, each action is an AA act mapping state realisations to differences between actions and the state.

7 Conclusion

In this paper, I develop a theory of BDP and belief updating that can be tested by observing choices over contingent menus. I conclude by providing a second interpretation of the model and discussing a few implications of the analysis.

The interpretation of the model in the main text is that the individual distorts signals in the direction of her BDP preferences and updates beliefs with Bayes rule via distorted signals. However, the model admits a second interpretation, closer in spirit to that of Epstein (2006). When observing the realisation of the contingent menu, the individual might revise her prior beliefs rather than distort the signal. The revised prior beliefs are updated according to the true signal, leading to posterior beliefs satisfying BDP. The conceptual distinction between these two interpretations relates to the supposed "irrationality" of motivated reasoning. Distorting the objective likelihood of a signal is arguably a mistake. However, revising a prior belief is not necessarily irrational. Since it is a subjective belief, there is no objective counterpart to qualify it as "wrong". Such an interpretation is possible in this model because it features subjective beliefs contrary to the objective lotteries framework in the past literature. Under this second interpretation, discussing the trade-off between accuracy and utility from beliefs is meaningless, as there is no "accurate" belief. Moreover, regardless of the subjective or objective nature of the belief constituting a benchmark, both the current and previous models feature individuals having both the "true" belief and the "distorted" belief in mind. If the individual can formulate the trade-off between accuracy and utility, she knows what is accurate, but how can she believe something else then? This trade-off is a critical component of the BDP literature (Bénabou & Tirole, 2016). If motivated reasoning is interpreted as a rational model of decision-making, the process leading to belief revision inconsistent with Bayes rule cannot be explained in terms of an accuracy-utility trade-off. There is a second element that favours the change in prior interpretation. The model requires independence, which I justified with indifference to probabilistic mixtures. Why should the individual distort the likelihood of a signal and not the probability mixture of two contingent menus? The change in the prior interpretation is consistent with the correct processing of information coming from objective lotteries of contingent menus in the model. I thus challenge the common wisdom that the accuracy-utility trade-off is a conceptually appealing tool for theories of motivated belief updating. A trade-off between material and belief-based utility, such as the one formulated in the first example in Section 2, seems more intuitive.

Adopting a model of BDP leading to non-Bayesian updating has implications for agreement theorems in the style of Aumann (1976). Two individuals with the same prior beliefs but different BDP have distinct posteriors beliefs, even if these are common knowledge. Individuals with the same BDP will instead have the same posterior beliefs. This does not necessarily follow from previous models, in which BDP and non-Bayesian updating are disjointed assumptions. This violation of Aumann makes BDP suitable candidates for explaining phenomena of polarisation and assortativity on the basis of preferences, as hinted in the application in Section 6.

References

Aliprantis, C. D., & Border, K. C. (2006). *Infinite dimensional analysis : A hitchhiker's guide* (3rd Ed. ed.). Berlin: Springer. 12, 18

Andreoni, J., Rao, J. M., & Trachtman, H. (2017). Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving. *Journal of Political Economy*, *125*(3), 625–653. 2

Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *Annals of mathematical statistics*, *34*(1), 199–205. 6, 11

Aumann, R. J. (1976). Agreeing to disagree. *The annals of statistics*, *4*(6), 1236–1239. 28

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *The American Economic Review*, *85*(5), 1337–1343. 2

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. 2

Battigalli, P., & Dufwenberg, M. (2022). Belief-dependent motivations and psychological game theory. *Journal of Economic Literature*, *60*(3), 833–882. 5, 6

Bénabou, R. (2015). The economics of motivated beliefs. *Revue d'économie politique*(5), 665–685. 2, 3, 24

Bénabou, R., & Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, *30*(3), 141–64. 2, 6, 24, 27

Brunnermeier, M. K., & Parker, J. A. (2005). Optimal expectations. *American Economic Review*, *95*(4), 1092–1118. 2, 3, 7, 23, 24

Caplin, A., & Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, *116*(1), 55–79. 2

Chan, T. Y., Liao, L., Martin, X., & Wang, Z. (2024). Avoiding Peer Information and Its Effects on Charity Crowdfunding: A Field Experiment. *Management Science*, *70*(4), 2272–2293. 2

Cobb-Clark, D. A., Dahmann, S. C., Kamhöfer, D. A., & Schildberg-Hörisch, H. (2022). The predictive power of self-control for life outcomes. *Journal of Economic Behavior & Organization*, *197*, 725–744. 5

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*, 67–80. 7

Daniel, K., & Hirshleifer, D. (2015). Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives*, *29*(4), 61–88. 2, 9

Derksen, L., Kerwin, J. T., Reynoso, N. O., & Sterck, O. (2024). Healthcare appointments as commitment devices. *The Economic Journal*, ueae077. 11

Dillenberger, D., & Raymond, C. (2020). Additive-belief-based preferences. *PIER Working Paper*. 6

Drobner, C., & Goerg, S. J. (2024). Motivated Belief Updating and Rationalization of Information. *Management Science*, *70*(7), 4583–4592. 2, 24

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–138. 2, 24

Eliaz, K., & Spiegler, R. (2006). Can anticipatory feelings explain anomalous choices of information sources? *Games and Economic Behavior*, *56*(1), 87–104. 4, 21

Epstein, L. G. (2006). An axiomatic model of non-Bayesian updating. *The Review of Economic Studies*, *73*(2), 413–436. 6, 23, 27

Epstein, L. G., & Kopylov, I. (2007). Cold feet. *Theoretical Economics*, *2*, 231–259. 6

Fishburn, P. C. (1970). *Utility theory for decision making*. New York: Wiley. 32, 34, 35, 36

Garrett, N., & Sharot, T. (2017). Optimistic update bias holds firm: Three tests of robustness following Shah et al. *Consciousness and cognition*, *50*, 12–22. 24

Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, *78*(1), 35–71. 2

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. *Journal of economic literature*, *55*(1), 96–135. 2, 6, 7, 9

Grossman, Z., & van der Weele, J. (2013). Self-image and strategic ignorance in moral dilemmas. *Working Paper*. 7

Gul, F., & Pesendorfer, W. (2001). Temptation and Self-Control. *Econometrica*, *69*(6), 1403–1435. 17, 22, 23, 39

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature climate change*, *2*(10), 732–735. 2

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615. 25, 26

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty*, *38*(2), 95–115. 2

Kopylov, I. (2009). Temptations in General Settings. *The B.E. Journal of Theoretical Economics*, *9*(1). 35, 40

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, *4*(4), 673–707. 2, 3, 23

Köszegi, B. (2010). Utility from anticipation and personal equilibrium. *Economic Theory*, *44*, 415–444. 2, 7

Liang, Y. (2017). Information-dependent expected utility. *Available at SSRN 2842714*. 4, 6, 21, 33, 34, 35, 37, 40

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2022). Managing Self-Confidence: Theory and Experimental Evidence. *Management Science*, *68*(11), 7793–7817. 2, 24

Niehaus, P. (2014). A theory of good intentions. *San Diego, CA: University of California and Cambridge, MA: NBER*, *111*. 7

Oster, E., Shoulson, I., & Dorsey, E. R. (2013). Optimal expectations and limited medical testing: Evidence from Huntington disease. *American Economic Review*, *103*(2), 804–830. 2

Ozdenoren, E. (2002). Completing the state space with subjective states. *Journal of Economic Theory*, *105*(2), 531–539. 6

Rommeswinkel, H., Chang, H.-C., & Hsu, W.-T. (2023). Preference for Knowledge. *Journal of Economic Theory*, *214*, 105737. 4, 6

Savage, L. J. (1972). *The foundations of statistics*. Dover Publications. 6, 24

Sicherman, N., Loewenstein, G., Seppi, D. J., & Utkus, S. P. (2016). Financial attention. *The Review of Financial Studies*, *29*(4), 863–897. 2

Siniscalchi, M. (2011). Dynamic choice under ambiguity: Dynamic choice under ambiguity. *Theoretical Economics*, *6*(3), 379–421. 7

Spiegler, R. (2008). On two points of view regarding revealed preference and behavioral economics. *The foundations of positive and normative economics: A handbook*, 95–115. 23

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. *American Economic Review*, *98*(5), 1829–1863. 2

Van der Weele, J. J. (2014). Inconvenient truths: Determinants of strategic ignorance in moral dilemmas. *Available at SSRN 2247288*. 7

von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior*. Princeton University Press. 4

Yariv, L. (2002). I'll see it when I believe it? A simple model of cognitive consistency. *Working Paper*. 24

APPENDIX

A PROOFS

Proof of Theorem 1. Necessity is omitted. I only prove sufficiency.

First, I study the properties of conditional rankings on menus induced by preferences over contingent menus. For each likelihood ℓ, the ranking $\succsim$ induces the conditional ranking $\succsim_\ell$ on the set of menus $\mathcal{M}$ defined as

$$M \succsim_\ell M' \text{ if } F \succsim F_{M \to M'} \text{ for some } F \text{ such that } \ell_{M,F} = \ell,$$

where F and $F_{M \to M'}$ coincide except for one realisation delivering M in the first when the second delivers $M' \notin \mathcal{M}_F$. Any two contingent menus F and $F_{M \to M'}$ satisfy II, as they coincide except when delivering M and M' inducing the same likelihood. By II Independence, any mixture of F, $F_{M \to M'}$ with a third contingent menu $F_{M \to M''}$ preserves their ranking. Therefore, for all likelihoods ℓ, the ranking $\succsim_\ell$ satisfies Independence on the set of menus $\mathcal{M}$. Due to Independence, the ranking between any two menus M and M' under $\succsim_\ell$ does not depend on the specific contingent menu F, as long as $\ell_{M,F} = \ell$. Therefore, Independence implies that "for some" in the definition of $\succsim_\ell$ is equivalent to "for all".

I now study properties of conditional preferences on particular singleton menus. The set of all contingent menus is $\mathcal{C}$. For any collection of likelihoods indexed by menus $\widehat{\ell} = \left(\widehat{\ell}_M\right)_{M \in \mathcal{M}}$, consider the set of contingent menus

$$\mathcal{C}_{\widehat{\ell}} := \left\{ F \in \mathcal{C} \;\middle|\; \ell_{M,F} = \widehat{\ell}_M \text{ and } M = \{f\} \text{ for one } f \in \Delta(X)^{\mathcal{S}} \text{ for all } M \in \mathcal{M}_F \right\}.$$

Contingent menus in any $\mathcal{C}_{\widehat{\ell}}$ only deliver distributions of singleton menus containing an act for each state s and the same likelihood $\widehat{\ell}_{\{f\}}$ for each menu realisation $\{f\}$ in their support. The set of such contingent menus is a set of AA acts $F : \mathcal{S} \to \Delta^\circ(M)$, each inducing the same likelihood at time 1 for each of the menu realisations in their support. The ranking $\succsim$ on each $\mathcal{C}_{\widehat{\ell}}$ satisfies Order, Continuity and Independence, and hence has a standard AA representation. By standard results (Fishburn, 1970, Theorem 13.1 pag. 176) preferences over contingent menus in $\mathcal{C}_{\widehat{\ell}}$ have the following representation

$$\mathscr{U}(F) = \sum_{\{f\}} \sum_s F_s(\{f\})\, U\left(f; \widehat{\ell}_{\{f\}}\right), \tag{10}$$

for all $F \in \mathcal{C}_{\widehat{\ell}}$, where $U\left(f;\widehat{\ell}_{\{f\}}\right)$ represents the conditional ranking $\succsim_{\widehat{\ell}_{\{f\}}}$ over singleton menus.

The next step is to employ Theorem 1 in Liang (2017) to show that Order, Continuity, II Independence and Nondegeneracy imply that preferences over menus depend on the likelihood these induce. Liang (2017) has a different setting, assumes the set of outcomes is infinite and has different axioms, so a few modifications of his proof are needed.

I now prove two preliminary lemmas that allow me to employ Liang's result. Recall that fsf' is an act equivalent to f in state s and to f' in all states $s' \neq s$.

Lemma 1. *Assume $\succsim$ satisfies Order, Continuity, and II Independence. Then, for any collection of likelihoods $\widehat{\ell}$, if $F \in \mathcal{C}_{\widehat{\ell}}$ and $\widehat{\ell}_{\{f'\}}(s) = 0$ for some f' and s, then $F \sim F_{\{f'\}\to\{fsf'\}}$ for all f.*

Proof. Preferences over contingent menus in each $\mathcal{C}_{\widehat{\ell}}$ are represented by Equation (10)

$$\mathscr{U}(F) = \sum_{\{f\}} \sum_{s} F_s(\{f\})\, U\left(f_s;\widehat{\ell}_{\{f\}}\right), \tag{10}$$

and therefore

$$\begin{aligned}\mathscr{U}\left(F_{\{f'\}\to\{fsf'\}}\right) = &\sum_{\{f''\}\neq\{fsf'\}} \sum_{s'} \left(F_{\{f'\}\to\{fsf'\}}\right)_{s'}(\{f''\})\, U\left(f''_s;\widehat{\ell}_{\{f\}}\right)\\ &+ \sum_{s'} \left(F_{\{f'\}\to\{fsf'\}}\right)_{s'}(\{fsf'\})\, U\left((fsf')_{s'};\widehat{\ell}_{\{f\}}\right).\end{aligned}$$

Due to the definition of fsf', it follows that the last term is equal to

$$\begin{aligned}&\sum_{s'\neq s} \left(F_{\{f'\}\to\{fsf'\}}\right)_{s'}(\{fsf'\})\, U\left(f'_{s'};\widehat{\ell}_{\{f\}}\right)\\ &+ \left(F_{\{f'\}\to\{fsf'\}}\right)_{s}(\{fsf'\})\, U\left(f_s;\widehat{\ell}_{\{f\}}\right).\end{aligned}$$

If $\widehat{\ell}_{\{f'\}}(s) = 0$ then $\widehat{\ell}_{\{fsf'\}}(s) = 0$ and $\left(F_{\{f'\}\to\{fsf'\}}\right)_s(\{fsf'\}) = 0$. The last term of the equation is thus equal to zero. Therefore, changing f' to f in state s leads to indifference. $\square$

Lemma 2. *The ranking $\succsim$ satisfies Order, Continuity, II Independence, Nondegeneracy if and only there exist a function $\mathscr{U}$ representing $\succsim$ that is mixture linear in contingent menus satisfying II.*

Proof. The set of finite probability distributions on $\mathcal{M}$ is a mixture space and each contingent menu $F : \mathcal{S} \to \Delta^{\circ}(\mathcal{M})$ is an AA act inducing only lotteries with finite support. Any two mixtures of two contingent menus satisfying II also satisfy II. Therefore, if $F \succsim F' \succsim F''$ and F, F'' satisfy II, there exists a unique λ such that $\lambda F + (1-\lambda) F'' \sim F'$ (Fishburn, 1970, Theorem 8.3 pag. 112).

Consider two contingent menus such that $F \succ F'$ and $F_s(\{x_s\}) = 1$ and $F'_s(\{x'_s\}) = 1$ for all s. These contingent menus induce a distinct singleton menu in each state, and therefore each of their realisation reveals the state. These two contingent menus exist by Nondegeneracy. For each contingent menu G such that $F \succ G \succ F'$, define $\mathscr{U}(G)$ to be equal to the unique λ such that $\lambda F + (1-\lambda) F' \sim G$.

For each contingent menu G satisfying II with F' and such that $G \succ F$, define $\mathscr{U}(G)$ to be $1/\lambda$ where λ is the unique number such that $\lambda G + (1-\lambda) F' \sim F$. For each contingent menu G satisfying II with F and such that $G \prec F'$, define $\mathscr{U}(G)$ to be $\lambda/(\lambda-1)$ where λ is the unique number such that $\lambda F + (1-\lambda) G \sim F'$.

Consider a contingent menu such that $G \prec F'$ but G and F do not satisfy II. This means that, for some s, $\{x_s\} \in \mathcal{M}_G$, the support of menus induced by the contingent menu G and $\ell_{\{x_s\},G}(s) \neq 1$. By Lemma 1, $F \sim F_{\{x_s\} \to \{fs'x_s\}}$ for all f and $s' \neq s$, since $\ell_{\{x_s\},F}(s') = 0$ for all $s' \neq s$. Because the support of any contingent menu is finite, there is always an act f such that the menu $\{fs'x_s\}$ is outside the support of G. Therefore, it is enough to define $\mathscr{U}(G)$ using the procedure above and mixing it with $F_{\{x_s\} \to \{fs'x_s\}}$. A similar construction works if $G \succ F$ but G and F' do not satisfy II.

By Proposition 1 in Liang (2017), the utility function $\mathscr{U}$ is well-defined, represents the ranking $\succsim$, and is linear in mixtures of contingent menus satisfying II. □

I now prove that Order, Continuity, II Independence, Nondegeneracy are equivalent to the following expected utility representation.

PROPOSITION 1. *The ranking $\succsim$ satisfies Order, Continuity, II Independence, Nondegeneracy if and only if it can be represented by*

$$\mathscr{U}(F) = \sum_{M} \sum_{s} p(s) F_s(M) \mathcal{U}(M; \ell_{M,F}) \tag{11}$$

for all contingent menus F, where $\mathcal{U} : \mathcal{M} \times \Delta(\mathcal{S}) \to \mathbb{R}$ is continuous and bounded on both $\mathcal{M}$ and $\Delta(\mathcal{S})$ and $p \in \Delta(\mathcal{S})$.

Proof. Since preferences over contingent menus $\succsim$ satisfy Order, II Independence, Continuity, and Nondegeneracy, by Lemmas 1, 2 and Theorem 1 in Liang (2017), the necessity and sufficiency of the representation in Equation (11) holds.

I now prove the continuity of $\mathcal{U}$ by contrapositive, I show that if it is not continuous, then Continuity does not hold. Suppose that $\mathcal{U}$ is not continuous at some point (M_0, ℓ_0) in $\mathcal{M} \times \Delta(\mathcal{S})$. Then, there exists $\varepsilon > 0$ such that for every $\delta > 0$, there is a point (M, ℓ) satisfying:

$$d\left(M, M_0\right) < \delta, \quad \|\ell - \ell_0\| < \delta, \quad \text{and} \quad |\,\mathcal{U}(M;\ell) - \mathcal{U}(M_0;\ell_0)\,| \geq \varepsilon,$$

where d is the Hausdorff metric. Consider a sequence $(M_n, \ell_n) \to (M_0, \ell_0)$. Construct a sequence of contingent menus F_n such that for each n, the menu M_n has $\sum_s p\left(s\right) F_{s,n}\left(M_n\right) > 0$, and the likelihood ℓ_{M_n,F_n}. Since the mapping from F_n to $\left(\sum_s p\left(s\right) F_{s,n}\left(M_n\right), \ell_{M_n,F_n}\right)$ is continuous, and $\mathscr{U}\left(F_n\right)$ depends on $\mathcal{U}(M;\ell_{M_n,F_n})$, the discontinuity in $\mathcal{U}$ at (M_0, ℓ_0) leads to a discontinuity in $\mathscr{U}\left(F_n\right)$ at the corresponding F_0. This contradicts Continuity, which requires $\mathscr{U}$ to be continuous in F (Fishburn, 1970, Theorem 3.5 p. 36). Therefore, $\mathcal{U}$ must be continuous on $\mathcal{M} \times \Delta\left(\mathcal{S}\right)$. Since $\mathcal{U}$ is continuous on a compact space, it is bounded. □

I now study the shape of $\mathcal{U}$. Because $\succsim$ satisfies Set-Betweenness, each $\succsim_\ell$ inherits the property

$$M \succsim_\ell M' \implies M \succsim_\ell M \cup M' \succsim_\ell M', \tag{12}$$

for all menus M, M'. By Order and Continuity, all $\succsim_\ell$ satisfy analogous properties. Since each ranking $\succsim_\ell$ satisfies Order, Continuity, Independence, and Set-Betweenness as expressed in Equation (12), by Theorem 1 in Kopylov (2009) it can be represented by

$$\mathcal{U}\left(M;\ell\right) = \max_{f \in M} \left\{ U\left(f;\ell\right) + V\left(f;\ell\right) - \max_{f' \in M} V\left(f';\ell\right) \right\} \tag{13}$$

for all menus M and likelihoods ℓ, where the functions $U\left(\cdot;\ell\right)$ and $V\left(\cdot;\ell\right)$ are continuous, bounded and linear in convex combinations of acts. Moreover, for all likelihoods ℓ, the functions $U\left(\cdot;\ell\right)$ and $V\left(\cdot;\ell\right)$ satisfy the mixture space axioms.

The state dependent version of AA theorem and their continuity imply that these have the following functional forms

$$U\left(f;\ell\right) = \sum_s u\left(f_s;\ell,s\right)\ ,\ V\left(f;\ell\right) = \sum_s v\left(f_s;\ell,s\right) \tag{14}$$

for all acts $f : \mathcal{S} \to \Delta(X)$. Each $u(\cdot;\ell,s)$ and $v(\cdot;\ell,s)$ is continuous, bounded, and linear in mixtures of lotteries over X.

The next step is to obtain the stronger state-independent subjective expected utility representation. Recall that $L \subseteq \Delta(X)$ denotes a menu of outcome lotteries and, for each state s and menus M, M', the definition of the menu $MsM' := \{fsf' \mid f \in M, f' \in M'\}$.

Lemma 3. *Assume $\succsim$ satisfies Order, State Independence, and Nondegeneracy. Then, for all menus L, L', M, likelihoods ℓ and states s, s' the ranking $\succsim_\ell$ satisfies*

$$LsM \succsim_\ell L'sM \;\Rightarrow\; Ls'M \succsim_\ell L's'M. \tag{15}$$

Moreover, for all menus M, likelihoods ℓ and states s, if $LsM \sim_\ell L'sM$ for all menus L and L', then $\ell(s) = 0$.

Proof. By definition of $\succsim_\ell$, the antecedent of condition (15) holds if $F \succsim F_{LsM \to L'sM}$ for some contingent menu F such that $\ell_{LsM,F} = \ell$. By State Independence, for each state s' and F, it holds that $F \succsim F_{Ls'M \to L's'M}$, which implies $Ls'M \succsim_\ell L's'M$.

If $\ell(s) \neq 0$ under the hypothesis of the second part of the Lemma, then Order and State Independence would imply that $L \sim_\ell L'$ for all L, L', contradicting Nondegeneracy. □

I now consider again preferences over contingent menus in $\mathcal{C}_{\widehat{\ell}}$. By Proposition 1, these are represented by the following expected utility function

$$\mathscr{U}(F) = \sum_{\{f\}} \sum_{s} p(s)\, F_s(\{f\})\, U\left(f; \widehat{\ell}_{\{f\}}\right). \tag{16}$$

By the first part of Lemma 3 and Nondegeneracy, each function u in the first part of Equation (14) has the subjective expected utility form (Fishburn, 1970, Theorem 13.2 pag. 177), and therefore Equation (16) can be rewritten as

$$\mathscr{U}(F) = \sum_{\{f\}} \sum_{s} \left(\sum_{s'} p(s')\, F_{s'}(\{f\}) \right) p_{\widehat{\ell}_{\{f\}}}(s)\, u\left(f_s; \widehat{\ell}_{\{f\}}\right).$$

The probability a menu $\{f\}$ realises is $\sum_{s'} p(s')\, F_{s'}(\{f\})$, while the probability state s and menu $\{f\}$ realise is $p(s)\, F_s(\{f\})$. By the chain rule, $p_{\widehat{\ell}_{\{f\}}}$ must be the Bayesian posterior of p and therefore for each menu $\{f\}$ and state s

$$p_{\widehat{\ell}_{\{f\}}}(s) := \frac{\widehat{\ell}_{\{f\}}(s)\, p(s)}{\sum_{s'} \widehat{\ell}_{\{f\}}(s')\, p(s')} = \frac{F_s(\{f\})\, p(s)}{\sum_{s'} F_{s'}(\{f\})\, p(s')}.$$

Without loss of generality, for each likelihood ℓ and state s

$$u(\cdot\,;\ell,s) = p_\ell(s)\,u(\cdot\,;\ell)\,, \tag{17}$$

where p_ℓ is the Bayesian update of the prior p under the likelihood ℓ. By Full Support the prior p has full support. By Equation (17), the first part of Equation (14) can be rewritten as

$$U(f;\ell) = \sum_s p_\ell(s)\,u(f_s;\ell)\,. \tag{18}$$

Lastly, I study the shape of V. For this purpose, I need to construct a specific class of contingent menus. The following is the set of all contingent menus inducing likelihood ℓ whenever menu M realises

$$\mathcal{C}_\ell^M := \{F \in \mathcal{C} \mid \ell_{M,F} = \ell\}\,.$$

Next, I define the set of preferred outcomes at likelihood ℓ

$$X_\ell := \left\{x \in \Delta(X) \,\middle|\, F \succsim F_{\{x\}\to\{x'\}} \text{ for all } x' \in \Delta(X) \text{ and some } F \in \mathcal{C}_\ell^{\{x\}}\right\}.$$

A generic element of X_ℓ is x_ℓ. Define contingent menu $\overline{F}$ so that $\overline{F}_s(\{x_s\}) = 1$ for all s, with $x_s \in \Delta(X)$. This is a contingent menu whose all menu realisations are singletons containing an outcome and revealing the state. Normalise $\mathscr{U}\left(\overline{F}\right) = 0$. For each event S, for each likelihoods $\ell \in \Delta(S)$, construct contingent menus $\overline{F}^\ell$ as follows. First, $\overline{F}^\ell$ coincides with $\overline{F}$ outside S, that is $\overline{F}_s^\ell(M) = \overline{F}_s(M)$ for each state $s \notin S$ and menu M. For each $s \in S$, two properties must hold:

1. $\overline{F}_s^\ell(\{x_\ell\}) = 1 - \overline{F}_s^\ell(\{x_s\})$;

2. $\ell_{\{x_\ell\},\overline{F}^\ell} = \ell$.

The contingent menu $\overline{F}^\ell$ always induces the likelihood ℓ when the singleton menu $\{x_\ell\}$ realises. By Theorem 1 in Liang (2017), for any contingent menu $\overline{F}^\ell$, there always exist a unique decomposition such that

$$\frac{1}{\mid\mathcal{S}\mid}\overline{F}^\ell + \left(\frac{\mid\mathcal{S}\mid -1}{\mid\mathcal{S}\mid}\right)\overline{F} = \lambda F^\ell + (1-\lambda)\,\overline{F},$$

where the contingent menu F^{ℓ} satisfies properties 1. and 2., plus $F_s^{\ell}(\{x_\ell\}) = \ell(s)$, and $\lambda = \frac{1}{|\mathcal{S}|}\sum_s \overline{F}_s^{\ell}(\{x_\ell\})$.

By Lemma 2 and because $\mathscr{U}\left(\overline{F}\right) = 0$, it follows that

$$\frac{1}{\mid \mathcal{S} \mid}\mathscr{U}\left(\overline{F}^{\ell}\right) + \left(\frac{\mid \mathcal{S} \mid -1}{\mid \mathcal{S} \mid}\right)\mathscr{U}\left(\overline{F}\right) = \lambda \mathscr{U}\left(F_s^{\ell}\right) + (1-\lambda)\,\mathscr{U}\left(\overline{F}\right)$$
$$\mathscr{U}\left(\overline{F}^{\ell}\right) = \sum_s F_s^{\ell}(\{x_\ell\})\,\mathscr{U}\left(F_s^{\ell}\right).$$

In the construction of Proposition 1, the function $\mathcal{U}(\{x_\ell\};\ell)$ is defined to be $\mathscr{U}\left(F_s^{\ell}\right)$. Moreover, since $\{x_\ell\}$ is a singleton menu whose only element is an outcome, by Equations (13) and (18), the function $\mathcal{U}(\{x_\ell\};\ell)$ represents the same ordering as $u(x_\ell;\ell)$. Therefore, observing choices among menus constructed as F_s^{ℓ} allows identifying the preferred likelihoods according to the ranking represented by u.

For each event S, define the likelihoods

$$\ell_S^* \in \left\{\ell \in \Delta(S)\,\middle|\, F^{\ell} \succsim F^{\ell'} \text{ for all } \ell' \in \Delta(S)\right\}. \tag{19}$$

These are the likelihoods in the statement of SRBL. Because $\mathscr{U}\left(F^{\ell}\right)$ represents the same ordering as $u(x_\ell;\ell)$, and because $x_\ell \in X_\ell$ for each ℓ, likelihoods satisfying condition (19) also satisfy Equation (5), for each event S

$$\ell_S^* \in \underset{\ell\in\Delta(S)}{\arg\max}\ \max_{x\in\Delta(X)}\ u(x;\ell). \tag{5}$$

The next step is to use SRBL to show that for each event S, for each likelihood $\ell \in \Delta(S)$

$$V(\cdot;\ell) = \alpha_\ell U(\cdot;\ell_S^*) + \beta_\ell \tag{20}$$

for some ℓ_S^* satisfying condition (19), with $\alpha_\ell \geq 0$ for each ℓ and for some number β_ℓ. That is, the temptation ranking at event S coincides with the ranking U when one of the preferred likelihoods at that event realises. I prove the claim by contrapositive, i.e., if $V(\cdot;\ell)$ represents any other ranking, then SRBL is violated.

Assume that, for an event S and a likelihood $\ell \in \Delta(S)$, the temptation ranking $V(\cdot;\ell)$ does not represent the same ranking as any $U(\cdot;\ell_S^*)$ over acts, for all ℓ_S^* satisfying condition (19). The implication is that there exist two acts f, f' such that $U(f;\ell_S^*) > U(f';\ell_S^*)$ and $V(f';\ell) > V(f;\ell)$. Consider the menus $M = \{f\}$ and $M' = \{f'\}$, and a likelihood ℓ such that the antecedent of SRBL holds, that is

$$\mathcal{F}_{M\cup M',\ell}\cap\mathcal{F}_{M\cup M',\ell_S^*}\neq\emptyset \text{ for at least one } \ell_S^*, \tag{21}$$

where one such ℓ always exists, since the set of likelihoods satisfying condition (19) is non-empty and taking ℓ from that set suffices. By Equations (11) and (21), there is a common maximal element of $U(\cdot;\ell)$ and $U(\cdot;\ell_S^*)$, for one ℓ_S^*, in $M\cup M'$, which by hypothesis must be f. However, the maximal element of $V(\cdot;\ell)$ is f', which, by Equation (11), implies that $F\succ F_{M\to M\cup M'}$, in violation of SRBL. Therefore, if Equation (20) is violated, then SRBL is violated.

By Equations (13), (18) and (20) it follows that for each menu M, event S and likelihood $\ell\in\Delta(S)$

$$\begin{aligned}\mathcal{U}(M;\ell)=&\max_{f\in M}\left\{\sum_s p_\ell(s)\,u(f_s;\ell)+\alpha_\ell\sum_s p_{\ell_S^*}(s)\,u(f_s;\ell_S^*)\right\}\\&-\max_{f'\in M}\alpha_\ell\sum_s p_{\ell_S^*}(s)\,u(f_s';\ell_S^*)\,,\end{aligned} \tag{22}$$

which, together with Equation (11) delivers the representation.

□

Proof of Corollary 1. I omit the necessity part of the statement. Suppose that both (u,p,α,ℓ^*) and (u',p',α',ℓ'^*) represent $\succsim$, where $\ell^*=(\ell_S^*)_{S\in 2^S}$ is a collection of likelihoods satisfying condition (19), one for each event. I first show the uniqueness properties of u and p. By AA subjective expected utility theorem, u' represents the same ranking as u if and only if, for all likelihoods ℓ, there exists $a_\ell,b_\ell\in\mathbb{R}_{++}\times\mathbb{R}$ such that

$$u'(\cdot;\ell)=a_\ell u(\cdot;\ell)+b_\ell \text{ and } p'_\ell=p_\ell. \tag{23}$$

For each likelihood $\ell\in\Delta(S)$, either

$$\alpha_\ell\neq 0 \text{ and } \ell_S^*\neq\ell \tag{24}$$

or not. The vector (u,p,α,ℓ^*) violates Equation (24) if and only if (u',p',α',ℓ'^*) also does. If this is the case, either $\alpha_\ell=0$, from which also $\alpha'_\ell=0$, or $\ell=\ell_S^*$, and then, by Equation (22), any couple of α_ℓ,α'_ℓ preserves the ordinal ranking. If Equation (24) holds, Theorem 4 in Gul & Pesendorfer (2001) implies that

$$V'(\cdot;\ell) = A_\ell V(\cdot;\ell) + B_{V,\ell} \tag{25}$$

$$U'(\cdot;\ell) = A_\ell U(\cdot;\ell) + B_{U,\ell} \tag{26}$$

where, by Theorem 1, for each event S and likelihood $\ell \in \Delta(S)$

$$V(f;\ell) = \alpha_\ell \sum_s p_{\ell_S^*}(s)\, u(f_s;\ell_S^*)$$

for all f and one ℓ_S^* satisfying condition (19). For each act f, event S and likelihood $\ell \in \Delta(S)$

$$\begin{aligned}
U'(f;\ell) &= \sum_s p_\ell(s)\, u'(f_s;\ell) \\
&= \sum_s p_\ell(s)\,(a_\ell u(f_s;\ell) + b_\ell) \\
&= a_\ell \sum_s p_\ell(s)\, u(f_s;\ell) + \sum_s p_\ell(s)\, b_\ell \\
&= a_\ell U(f;\ell) + B_{U,\ell},
\end{aligned}$$

from which $A_\ell = a_\ell$ for all likelihoods ℓ. By the uniqueness result in Theorem 1 of Liang (2017), the function $\mathcal{U}'$ represents the same ranking as $\mathcal{U}$ if and only if there exist $(a,b,c) \in \mathbb{R}_{++} \times \mathbb{R} \times \mathbb{R}^S$ such that for all likelihoods ℓ

$$\mathcal{U}'(\cdot;\ell) = a\mathcal{U}(\cdot;\ell) + b - \sum_s c(s)\, p_\ell(s)\,.$$

Define the set of all contingent menus only containing singleton menus in their support

$$\overline{\mathcal{C}} := \left\{ F \in \mathcal{C} \,\middle|\, M = \{f\} \text{ for some } f \in \Delta(X)^S \text{ for all } M \in \mathcal{M}_F \right\}.$$

By Theorem 1, preferences over $\overline{\mathcal{C}}$ are represented by

$$\mathscr{U}(F) = \sum_{\{f\}} \sum_s p(s)\, F_s(\{f\})\, U\left(f;\ell_{\{f\},F}\right)$$

for all $F \in \overline{\mathcal{C}}$. Therefore, U is defined as $\mathcal{U}$ for singleton menus, and it inherits its uniqueness properties (Kopylov, 2009). In the present setting

$$\mathcal{U}'(\{f\};\ell) = a_\ell U(f;\ell) + B_{U,\ell}$$
$$= a_\ell \mathcal{U}(\{f\};\ell) + B_{U,\ell},$$

from which $a_\ell = a$ and $B_{U,\ell} = b - \sum_s c(s)\, p_\ell(s)$ for all ℓ. Because of the functional form of $\mathcal{U}$, the function $\mathcal{U}'$ only represents the same ranking if $c = 0$. In fact, for each event S and likelihood $\ell \in \Delta(S)$, substitution into the representation delivers

$$\begin{aligned}
\mathcal{U}'(M;\ell) &= \max_{f\in M}\left\{U'(f;\ell) + V'(f;\ell) - \max_{f\in M} V'(f;\ell)\right\} \\
&= \max_{f'\in M}\left\{aU(f;\ell) + b - \sum_s c(s)\, p_\ell(s)\right. \\
&+\alpha_\ell\left(aU(f;\ell^*_S) + b - \sum_s c(s)\, p_{\ell^*_S}(s)\right) \\
&\left.-\max_{f'\in M}\alpha_\ell\left(aU(f';\ell^*_S) + b - \sum_s c(s)\, p_{\ell^*_S}(s)\right)\right\}.
\end{aligned}$$

When taking expectations of $\mathcal{U}$, averaging the term $\sum_s c(s)\, p_\ell(s)$ makes it constant and equal to the prior for any likelihood, allowing to preserve the ranking.[16] However, the same does not hold for $\sum_s c(s)\, p_{\ell^*_S}(s)$. Therefore, to preserve ordinal equivalence, c must be null, otherwise the expression $\sum_s c(s)\, p_{\ell^*_S}(s)$ does not average to the prior and U' does not represent the same ranking as U. Moreover, by Equation (23) it follows that $b_\ell = b$ for each ℓ.

Next, I derive the uniqueness of α_ℓ for each likelihood ℓ. For each event S and likelihood $\ell \in \Delta(S)$, substitution of u in the expression of V delivers

[16] Consider the representation in Theorem 1. The algebra is as follows:

$$\begin{aligned}
\sum_M\sum_s p(s)\, F_s(M) \sum_{s'} c(s')\, p_{\ell_{M,F}}(s') &= \sum_M\sum_s p(s)\, F_s(M)\sum_{s'} c(s')\, \frac{F_{s'}(M)\, p(s')}{\sum_{s''} F_{s''}(M)\, p(s'')} \\
&= \sum_M\sum_s c(s)\, F_s(M)\, p(s) \\
&= c\,.
\end{aligned}$$

$$
\begin{aligned}
V'(f,\ell) &= \alpha'_\ell \sum_s p_{\ell^*_S}(s)\left(au\left(f_s;\ell^*_S\right)+b\right) \\
&= a\alpha_\ell \sum_s p_{\ell^*_S}(s)\, u\left(f_s;\ell^*_S\right)+B_{V,\ell} \\
&= aV(f,\ell)+B_{V,\ell}
\end{aligned}
$$

for all f, where the last equality follows from Equation (25) and the fact that $A_\ell = a$ for all ℓ. Therefore, $\alpha_\ell = \alpha'_\ell$ for each ℓ to preserve the ordinal ranking.

Lastly, I show that, for each event S, if the likelihood ℓ'^*_S part of ℓ'^* represents the same preferences over contingent menus as ℓ^*_S, then $\ell'^*_S(s) = \ell^*_S(s)$ for all states $s \in S$. I prove it by contrapositive, if this is not the case, then SRBL is violated.

First, I show that any two ℓ^*_S and ℓ'^*_S must induce the same posterior beliefs. Fix an event S and assume that both ℓ'^*_S and ℓ^*_S satisfy Equation (19) and that $p_{\ell'^*_S}(s) \neq p_{\ell^*_S}(s)$ for some s. Then, $U(\cdot,\ell'^*_S)$ does not represent the same ordering over acts of $U(\cdot,\ell^*_S)$. Assume this is not the case and these represent the same ordering over acts. Then, the representation of the ranking over constant acts $u(\cdot;\ell'^*_S)$ must be an affine transformation of $u(\cdot;\ell^*_S)$. However, since by hypothesis $p_{\ell'^*_S}(s) \neq p_{\ell^*_S}(s)$ for some s, and by Equation (18), $U(\cdot;\ell'^*_S)$ does not represent the same ordering over acts of $U(\cdot;\ell^*_S)$, which is absurd.

By Theorem 1, preferences over menus at ℓ^*_S can be represented by the following

$$
\mathcal{U}(M;\ell^*_S) = \max_{f\in M}\left\{U(f;\ell^*_S)+\alpha_{\ell^*_S}U(f;\ell'^*_S)-\max_{f'\in M}\alpha_{\ell^*_S}U(f';\ell'^*_S)\right\},
$$

for each menu M, since ℓ'^*_S satisfies condition (19). For any menu M, the antecedent of SRBL holds, as ℓ^*_S satisfies Equation (19) and therefore trivially

$$
\mathcal{F}_{M\cup M',\ell^*_S} \cap \mathcal{F}_{M\cup M',\ell''^*_S} \neq \emptyset \text{ for at least one } \ell''^*_S \text{ satisfying Equation (19).}
$$

However, if $U(\cdot;\ell'^*_S)$ and $U(\cdot;\ell^*_S)$ do not represent the same ordering over acts, its consequent will not hold in general. Consider two acts f, f' such that $U(f;\ell^*_S) > U(f';\ell^*_S)$ and $U(f';\ell'^*_S) > U(f;\ell'^*_S)$ and construct menus $M = \{f\}$ and $M' = \{f'\}$. By Equation (11), $F \succ F_{M\to M\cup M'}$ for any F such that $\ell_{M,F} = \ell^*_S$, in violation of SRBL. Therefore, if $p_{\ell'^*_S}(s) \neq p_{\ell^*_S}(s)$ for some s, then SRBL is violated.

Fixing a prior p, there is a one to one relationship between likelihood and posterior. For each likelihood ℓ and state s

$$\ell(s) = \frac{\frac{p_\ell(s)}{p(s)}}{\sum_{s'} \frac{p_\ell(s')}{p(s')}}.$$

By Equation (23), the prior p representing preferences is unique. Therefore, since the posteriors satisfy $p_{\ell'^*_S}(s) = p_{\ell^*_S}(s)$ for each s, it follows that $\ell'^*_S(s) = \ell^*_S(s)$ for each s.

□

B Construction of Best Likelihoods

I here provide an example of the construction of contingent menus that allows to identify the preferred likelihoods in Theorem 1. I start from these two contingent menus, F and $\overline{F}$, where the utility of the second is normalised to 0:

$$\begin{array}{cc} F & \overline{F} \\ \begin{bmatrix} \{x_\ell\} & (0.1, 0.4) \\ \{y\} & (0.9, 0) \\ \{z\} & (0, 0.6) \end{bmatrix} & \begin{bmatrix} \{x_{s_1}\} & (1, 0) \\ \{x_{s_2}\} & (0, 1) \end{bmatrix}. \end{array}$$

The aim is to represent a combination of these two contingent menus as a combination of three contingent menus, one for each menu in the support of F. The combination will be such that a part of it averages to give 0 utility and the rest gives the utility of each menu in the support of F at the likelihood it induces in that menu, preserving their realisation probability. Each of the three contingent menus is then defined as the utility of choosing from their corresponding menu in F at their likelihood.

Each of these contingent menus must contain one menu realisation from F at the same likelihood. So, for each menu in the support of F, I change the probability of its realisation in each state until it coincides with its normalised likelihood. Then , I fill the rest of the contingent menu with elements from $\overline{F}$. As an example, for $\{x_\ell\}$

$$\begin{bmatrix} \{x_\ell\} & (0.2, 0.8) \\ \{x_{s_1}\} & (0.8, 0) \\ \{x_{s_2}\} & (0, 0.2) \end{bmatrix}.$$

All three of them are as follows:

$$\begin{bmatrix} \{x_\ell\} & (0.2, 0.8) \\ \{x_{s_1}\} & (0.8, 0) \\ \{x_{s_2}\} & (0, 0.2) \end{bmatrix} \qquad \begin{bmatrix} \{y\} & (1, 0) \\ \{x_{s_2}\} & (0, 1) \end{bmatrix} \qquad \begin{bmatrix} \{z\} & (0, 1) \\ \{x_{s_1}\} & (1, 0) \end{bmatrix}.$$

I must construct a linear combination of these that coincides with a linear combination of the original contingent menus F and $\overline{F}$. In F, conditional on a state, a menu M realises with probability $F_s(M)$. In the three new contingent menus, conditional on a state, the probability that a menu M in the support of F realises is its normalised likelihood, namely $\frac{F_s(M)}{\sum_{s'} F_{s'}(M)}$. Therefore, to make the conditional probability of realisation coincides, the weight on each new contingent menu must be $\sum_{s'} F_{s'}(M)$, to cancel the denominator. However, summing for each menu yields

$$\sum_{M \in \mathcal{M}_F} \sum_{s'} F_{s'}(M) = \mid \mathcal{S} \mid,$$

which is greater than 1. Therefore, the weight on each new contingent menu should be $\frac{\sum_{s'} F_s'(M)}{|\mathcal{S}|}$, which results in the following linear combination:

$$0.25 \begin{bmatrix} \{x_\ell\} & (0.2, 0.8) \\ \{x_{s_1}\} & (0.8, 0) \\ \{x_{s_2}\} & (0, 0.2) \end{bmatrix} + 0.45 \begin{bmatrix} \{y\} & (1, 0) \\ \{x_{s_2}\} & (0, 1) \end{bmatrix} + 0.3 \begin{bmatrix} \{z\} & (0, 1) \\ \{x_{s_1}\} & (1, 0) \end{bmatrix}.$$

Since the conditional probability of each menu has been divided by the number of states, the probability of realisation of F in combination with $\overline{F}$ should be $\frac{1}{|\mathcal{S}|}$, to make conditional probabilities coincide

$$\frac{1}{2} \begin{bmatrix} \{x_\ell\} & (0.1, 0.4) \\ \{y\} & (0.9, 0) \\ \{z\} & (0, 0.6) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \{x_{s_1}\} & (1, 0) \\ \{x_{s_2}\} & (0, 1) \end{bmatrix}$$

$$= 0.25 \begin{bmatrix} \{x_\ell\} & (0.2, 0.8) \\ \{x_{s_1}\} & (0.8, 0) \\ \{x_{s_2}\} & (0, 0.2) \end{bmatrix} + 0.45 \begin{bmatrix} \{y\} & (1, 0) \\ \{x_{s_2}\} & (0, 1) \end{bmatrix} + 0.3 \begin{bmatrix} \{z\} & (0, 1) \\ \{x_{s_1}\} & (1, 0) \end{bmatrix}.$$

The conditional probability any menu in $\overline{F}$ realising also coincides in both linear combination, as it is

$$
\begin{aligned}
\sum_{M \in \mathcal{M}_F} \frac{\sum_{s'} F_{s'}(M)}{\mid \mathcal{S} \mid} \left(1 - \frac{F_s(M)}{\sum_{s''} F_{s''}(M)}\right) &= \frac{1}{\mid \mathcal{S} \mid} \sum_{M \in \mathcal{M}_F} \left(\sum_{s''} F_{s''}(M) - F_s(M) \right) \\
&= \frac{1}{\mid \mathcal{S} \mid} \left(\sum_{M \in \mathcal{M}_F} \sum_{s''} F_{s''}(M) - \sum_{M \in \mathcal{M}_F} F_s(M) \right) \\
&= \frac{1}{\mid \mathcal{S} \mid} (\mid \mathcal{S} \mid -1) \\
&= \frac{\mid \mathcal{S} \mid -1}{\mid \mathcal{S} \mid}.
\end{aligned}
$$

C Notation

Symbol [elements]	**Name**	**Mathematical object**
X	outcomes	compact metric set
$\Delta(X)$ $[x, y]$	(lotteries over) outcomes	compact metric set
$\mathcal{S}$ $[s, s']$	states	finite set
$S \subseteq \mathcal{S}$	events	finite set
f, f'	acts	functions $\mathcal{S} \to \Delta(X)$
$\mathcal{M}$ $[M, M']$	menus	compact metric set
$\Delta^{\circ}(\mathcal{M})$	finite lotteries over menus	compact metric set
$\mathcal{C}$ $[F, F']$	contingent menus	functions $\mathcal{S} \to \Delta^{\circ}(\mathcal{M})$
$\succsim$	preference	subset of $\mathcal{C} \times \mathcal{C}$
ℓ	likelihoods	probability distribution over $\mathcal{S}$
p	prior beliefs	probability distribution over $\mathcal{S}$
p_ℓ	posterior beliefs	probability distribution over $\mathcal{S}$
u	utility functions	functions $\Delta(X) \times \Delta(\mathcal{S}) \to \mathbb{R}$
$\mathcal{U}$	utility functions	functions $\mathcal{M} \times \Delta(\mathcal{S}) \to \mathbb{R}$
$\mathscr{U}$	utility functions	functions $\mathcal{C} \to \mathbb{R}$

Table 3: Symbols, their names, and corresponding mathematical objects.