COLLEGIUM OF ECONOMIC ANALYSIS
WORKING PAPER SERIES

Learning own preferences through consumption

Marek Kapera

SGH KAE Working Papers Series Number: 2022/074 April 2022

Learning own preferences through consumption

Marek Kapera*

Abstract

This paper provides theoretical foundations for preference discovery theory. We propose to relax the assumption that the consumer has perfect knowledge of their own preferences, so that the consumer knows only the subjective probability of those alternatives being in any given relation, which is conditional on the information available to the consumer. To achieve that, we construct probabilistic measures on the space of all permissible preference relations and consider the consumer to be equipped with one such measure, instead of a preference relation. These measures are intrinsically linked by construction to the information structure available to the consumer and allow for indirect learning. We visualize how these measures correspond to the choices of the consumer, we consider three distinct decision procedures. These procedures formalize how under different assumptions regarding the underlying probability measure, the consumer guesses their own tastes. Finally, we use these measures to define value of the information provided by the consumption of a chosen alternative and study the properties of the preference ranking induced by it.

Keywords: Taste uncertainty, Preference discovery, Learning through consumption, Conditional preferences, Experimental preferences

JEL classification: D11, D83, D91

*Institute of Economics, Polish Academy of Sciences, 00-330 Warsaw, Nowy Swiat 72, Poland, kaperamar@gmail.com, +48 609 686 255

1 Introduction

The most usual approach taken by economists to decision theory assumes what Simon (1976) calls the substantive rationality of individuals. This describes consumer choices as the outcome of utility maximising behaviour, that is driven by stable and well-defined consumer preferences. The approach differentiates the economic perspective on decision theory from the psychological one, which assumes procedural rationality. Substantive rationality is a stronger condition to satisfy, than procedural rationality, as it demands that rational behaviour be appropriate for achieving a certain goal under given limits and constraints, whereas procedural rationality sees behaviour as rational if only it is the outcome of appropriate mental deliberation.

Although the classical approach of economists (e.g. Debreu 1959) has substantial descriptive and predictive power, it comes at the price of multiple behavioral paradoxes (e.g. Kahneman and Tversky 1979, DellaVigna 2009, Lerner et al. 2015). While the examples of behaviour that many of those paradoxes give can be seen as merely violating the axioms of certain economic theories, some of the paradoxes, including contingent valuations (e.g. Schkade and Payne 1994, Loomis 2011), the endowment effect (e.g. Kahneman et al. 1991, Loewenstein and Issacharoff 1994, Bordalo et al. 2012a) and asymmetric dominance (Ariely and Wallsten 1995, Simonson and Tversky 1992) can be seen as challenges to the overall notion of substantive rationality. Especially troubling for economic theory is the preference reversal paradox (e.g. Lichtenstein and Slovic 1971, 1973, Bleichrodt and Prades 2009, Trautmann et al. 2011) which has led many psychological theories to reject the existence of preferences altogether. As Grether and Plott (1979) note, preference reversal can be seen as a paradox that undermine not just one particular economic theory, but rather the existence of any maximising behaviour at all. Many theories follow up on this rejection, giving rise to the notion of preference construction (e.g. Lichtenstein and Slovic 2006, Warren and McGraw 2011).

Theories of preference construction see choice as an outcome of procedures that the consumer follows when faced with a decision. Crucially, choice is not assumed to be procedure-invariant, meaning that different procedures give rise to different preferences (e.g. Tversky et al. 1988). People are assumed to use a wide variety of procedures, which they often develop on the spot when they are faced with a given choice, and even change during the task if the procedure leads to an unsatisfactory result or does not solve the problem at all (e.g. Coupey 1994, Slovic 1995, Bettman et al. 1998, Simon et al. 2008). Theories of preference construction often assume there to be different decision modes, like pairwise comparisons and separate evaluations (e.g. Schkade and Johnson 1989, Nowlis and Simonson 1997) and this is how psychological theory usually explains the preference reversal paradox. The procedures proposed include the search for a dominance structure, attribute weighting, salient choice and constraint satisfaction (e.g. Montgomery 1998, Tversky et al. 1988, Simon et al. 2004, Bordalo et al. 2012b, Glöckner et al. 2010).

Many theories of the substantively rational consumer have also been shown to be able to incorporate preference reversal. Those theories include regret theory (Loomes and Sudgen 1982), reference dependence (Sugden 2003), preference discovery costs (Wilson 2018) and contextual deliberations (Guo 2021). Of special importance to us however is the preference discovery theory of Plott (1996), which occupies the middle ground between economics and psychology. This theory retains the substantive rationality of the consumer but changes the interpretation of it, seeing it more as a process than as a state (e.g. Rizzo and Whitman 2018). At the same time it maintains the irrational character of paradoxical behaviour rather than accommodating it. It is a very intuitive theory that effectively uses Aristotle's concept of tabula rasa, or the lack of any innate knowledge, as its justification. Plott (1996) suggests that preference discovery is the theory that both economists and psychologists usually believe, but that they hardly ever state.

The theory proposed in Plott (1996) is based around the observation, that many of the paradoxes of choice, including preference reversal and the endowment effect (e.g. Cox and Grether 1994, Plott 1996), seem to be less prevalent in repeated experiments, with incentives. This observation led Plott (1996) to propose a three-stage theory of how the consumer discovers their own preferences. In the first stage, consumer choices are usually chaotic and irrational. It is during this stage that the consumer explores the set of choice alternatives available to them, learning by trial and error how much satisfaction they can obtain from each choice. In the second stage, the consumer's choices begin to stabilise, as the consumer starts to become aware of their own preferences. This is reflected by those choices getting closer and closer to what substantive rationality would imply. The only thing left at this point is for the behavior of other agents to be recognised as rational, and this is reflected in the continued prevalence of paradoxes in a setting that incorporates interactions with others. The recognition is finally achieved in the final stage, when the individual finally acts as a substantively rational consumer.

Rizzo and Whitman (2018) describe substantive rationality as a process from the point of view of preference discovery, rather than a given state. Preference discovery asserts that while well-defined and stable preferences exist, the choices observed can nevertheless be unstable if the individual is unaware of their own preferences. This differentiates preference discovery from the psychological viewpoint that preferences do not exist, and also from the dominant approach in economics, which is best described by the assertion of Stigler and Becker (1977) that we can always assume that preferences are given and stable.

Several studies (e.g. Braga and Starmer 2005, Braga et al. 2009, Bruni and Sudgen 2007) point out that although it can be seen as arguing in favour of the substantive rationality of the agent, preference discovery cannot be treated as a blanket defence of classical economic theory. Especially notable from our perspective is the observation of Bruni and Sudgen (2007) that we cannot simply assume that all people operating in markets have finished the process of preference discovery and are fully informed of their own tastes. This opinion finds support in experimental results (e.g. Kingsley and Brown 2010, Delaney et al. 2019) that show that even though individuals discover their preferences to a certain degree and this leads them to make more consistent choices, they tend not to discover their preferences fully, especially when the choices are important or the options are easy to compare (e.g. Hoeffler and Ariely 1999).

We agree with those sentiments, but rather than rejecting the preference discovery hypothesis because it fails as a defence of substantive rationality, we believe that it is precisely because of this failure that we should take preference discovery seriously and treat it as a standalone theory of consumer choice under imperfect knowledge of self. Intuitively, there are many instances in which preference discovery can contribute greatly to our understanding of consumer behaviour, and this remains the case even if consumers have tended to achieve a perfect discovery of their preferences in an experimental setting. Those instances can include the introduction of a new product; the possible inherent preference of the consumer for products they have not previously consumed, such as new movies; the acquisition of knowledge that causes a change that is hard to reverse, such as the discovery of preferences for addictive substances; or the boundaries placed on unlimited experimentation, as is usually the case with preferences for romantic partners. Preference discovery also seems like a natural theory to use for studying the experimental consumption of the individual, as learning about their own preferences through consumption gives that individual the motivation for such behaviour.

There is a lot of promising empirical research on preference discovery. There are studies that show that preference discovery can account for preference reversal (e.g. Cox and Grether 1996, Plott 1996, Butler and Loomes 2007), explain the WTP/WTA disparity (e.g. Plott and Zeiler 2005, Engelmann and Hollard 2010, Humphrey et al. 2017) and the order effects in stated preference studies (e.g. Day et al. 2012, Carlsson et al. 2012), while there are others that conclude that market experience is crucial in stabilising consumer choices (e.g. Kingsley and Brown 2010, Czajkowski et al. 2015). The results obtained by van de Kuilen (2007) suggest that preference discovery can even account for behavioural effects such as probability weighting, as the elicited probability weighting function converges significantly towards linearity when the respondents are asked to make repeated choices. At the same time, there are hardly any theoretical studies on the process of preference discovery, as Delaney et al. (2014) propose the only theoretical model of preference discovery, to the best of our knowledge.

Delaney et al. (2014) study the consumer with undiscovered preferences. They consider a very restrictive setting with a finite number of choice objects, and the menus of choices that are given exogenously in all time periods. Among the assumptions of the model are firstly that the consumer has exogenously given preferences that are well-behaved in all time periods, meaning they are complete, transitive and reflexive, and these preferences correspond to the consumer's guess about the real ranking of alternatives; secondly that the consumption of any given alternative does not change the ranking between any two different alternatives, meaning that indirect leaning is explicitly prohibited. Indirect learning means that consuming an alternative informs the consumer not only about its ranking relative to the other alternatives experienced, but also about the probability that one of the two unrelated alternatives is preferred over the other. As a result of those assumptions, the model of Delaney et al. (2014) only allows for the study of how preference discovery depends on the menus available to the consumer at all time periods and cannot be a satisfactory model of preference discovery.

That said, taste uncertainty has been found in the economic literature since at least the contribution of Kreps (1979). He showed that weak axioms on preferences are sufficient to represent introspective uncertainty using a subjective set of states. The result is that consumers can have a coherent, albeit not unique, subjective state space, without it being given exogenously. This has been achieved by a change of the domain over which the preferences are defined, meaning by defining the preferences for menu and consumption pairs rather than just for consumption, with the additional assumption of a preference for flexibility. The preference for flexibility states that the consumer should prefer a menu that is larger with respect to inclusion. One potential justification for this axiom is the existence of taste shocks that might affect future preferences, which then make a larger menu a safer choice. In a later reformulation by Dekel et al. (2001), the uniqueness of this subjective state space is obtained by another change of domain, this time to include lotteries over future action in a choice set. In both of these models, though, the resolution of uncertainty occurs irrespective of consumption choices. Cooke (2017) and Piermont et al. (2016) provide two extensions of the model that condition learning of the subjective state space on consumption and those two articles constitute the contemporary work that is most relevant to our study.

Both Cooke (2017) and Piermont et al. (2016) consider the consumer with uncertain tastes that exhibits the preference for flexibility. The demand for flexibility in those two models is driven by the desire to have more options, as any new information that is acquired can impact the choices from future menus. The main difference between the two models is in their setup: Piermont et al. (2016) consider ordinal preferences and infinite horizon of consumption, but the model of Cooke (2017) features two stages of consumption and cardinal preferences. Both models obtain the unique representation that is similar to the one in Kreps (1979) by a further change of domain: Cooke (2017) assumes preferences to be defined over the pairs of current consumption and future menu, and in Piermont et al. (2016) the objects of choice are infinite-horizon choice problems, defined as the paths of future consumption conditional on the information obtained. Indirect learning is permitted by both models, but only Cooke (2017) considers it in any detail.

Other notable work on taste uncertainty is the model proposed by Loomes et al. (2009). It is based on the earlier work on the reference dependent prospect theory by Sugden (2003), with the addition of a finite set of states of the world to represent taste uncertainty. The objects of choice in this setting are Savage's (1954) acts, and the reference act is chosen explicitly. The authors use the representation obtained to explain trade asymmetries and study how this effect changes with the change in taste uncertainty. However, Loomes et al. (2009) does not consider the connection between consumption and the resolution of taste uncertainty. There is no learning or any explicit modelling of preference discovery in this model.

In this article we propose theoretical foundations for preference discovery. We only consider the first stage of preference discovery from the original theory of Plott (1996), as we do not look at social interactions and restrict our attention only to the case of the consumer who is in the process of exploring the available alternatives. Moreover, we focus explicitly on a static setting, meaning we do not consider the questions about the evolution of preferences after choice or the paths of consumption generated by consumer choices. We only study how the consumer perceives their preferences given the information currently available to them. This perspective seems to us to be of particular importance, as most of the theoretical research on preference uncertainty (e.g. Delaney et al. 2014, Piermont et al. 2016, Cooke 2017) is focused on dynamic aspects of preference discovery.

We consider a consumer with limited self-knowledge, meaning a consumer who does not know their "real" preference relation, which is partially revealed by consumption. We incorporate both direct and indirect learning, as consumption not only reveals truthfully the preference rankings of the consumed alternatives, but also impacts on the perception of other alternative choices. This impact is introduced using the notion of similarity between alternative choices, meaning the consumer uses a similarity metric to extrapolate their preferences from the known alternatives to the unknown ones.

More specifically, we model the consumer's perception of their own preferences using probability measures on the space of all preference relations that satisfy certain rationality conditions, and the perception of the consumer's own preferences after they have consumed some limited subset of alternatives is obtained as a conditional probability measure. This probability is interpreted as the strength of the belief that any given preference relation is identical to the unknown, real preferences of the consumer. As those probability measures fully represent the consumer, we propose that the consumer be considered as equipped with such a probabilistic measure, rather than just a single preference relation. From this subjective probability, we identify and study two natural motivations that affect consumer choices during the preference discovery. These are firstly experimental motivation, which describes the value to the consumer of information about their own preferences as revealed by a given choice, and secondly immediate consumption-oriented motivation, which describes the current beliefs of the consumer about their real preferences, and focuses on the expected immediate utility from a given choice.

The starkest difference between our paper and contemporary works concerns the choice of the issues we study and the philosophy behind doing so. What we do is philosophically very close to various theories of preference construction (e.g. Lichtenstein and Slovic 2006), with the assumption that the real preference relation that informs the consumer about their own preferences is the only easily discernible red line that we do not cross. Even though this assumption is mostly philosophical and not binding in the static case, it is reflected throughout our article, in the assumption for example that consumer perception is path independent, meaning that only the subset of known alternatives matters, not the order in which they were consumed. However, we do not exclude the possibility that consumer choices are procedure dependent as we do not commit to any particular position on this and we do not propose any specific function for representing consumer choices. We do, however, consider multiple decision modes and how they correspond to different preferences, and we give examples of where evaluation in a given mode might be seen as reasonable. These examples might be interpreted either as legitimate decision procedures, or as specific examples of a more general choice function, but our goal is simply to describe different factors in consumer perception.

As a consequence we strongly diverge from contemporary articles in our modelling approach. Piermont et al. (2016), Cooke (2017) and Loomes et al. (2009) all assume that the consumer is equipped with exogenously given preferences over final objects of choice, that depending on the model in question can be menu and consumption bundles, infinite horizon choice problems or Savage's acts. These preferences exogenously determine the resolution of the trade-off between the immediate satisfaction from consumption and experimentation. We do not make that assumption, and instead we equip the consumer with a probability measure that represents how the consumer anticipates both learning and immediate satisfaction from consumption to be, but that does not assume a resolution of the conflict between those two motivations. Moreover we do not assume the preference for flexibility and do not use reference dependence or prospect theory.

Additionally we take a different approach to learning from consumption, that we model explicitly. In contrast to Piermont et al. (2016) and Cooke (2017) we assume an objective state space, that we define to be the state space to be a space of all preference relations that agree with the ranking of already consumed alternatives. Moreover not only like these articles we allow for indirect learning, we additionally use the notion of similarity between choices to provide additional structure that governs this behaviour. The idea of similarity as a tool that allows for indirect learning from experience is very much present in the literature, with Huang et al. (2014) also employing it in the context of incomplete preferences and more relevantly in the case-based decision theory proposed by Gilboa and Schmeidler (2001). This work is notable as these authors also consider the problem of evaluating future choices from past experiences. However, in case-based decision theory learning is not based on consumption, but rather on the vague memories of the decision maker. Moreover, the decision maker it considers is not substantively rational and the experimental behaviour is not motivated by the potential for higher utility from future choices, but rather by a general lack of satisfaction with the outcomes currently obtained, which is represented by the marginal value of the outcome that is deemed acceptable for the consumer.

The structure of the article is as follows. We start in section 2 with some preliminary definitions that form the conceptual framework for our work. Section 3 is mostly technical and occupies itself with the construction of a topology on the space of all the preference relations that the consumer might have, and of Borel measures on this space. We interpret those measures as representing the perception of the consumer about their own preferences, conditional on the information available. Sections 4 and 5 consider two major sources of motivation for the consumer, with immediate satisfaction from consumption in section 4 and discovery of the consumer's own preferences in section 5. Finally in section 6 we summarise everything we have done and point towards some possible areas for further research.

2 Elementary definitions

Let $\mathcal{B}$ be the space of choice objects. We assume that $\mathcal{B}$ comes equipped with a metric $d : \mathcal{B} \times \mathcal{B} \to \mathbb{R}_+$, together with topology on $\mathcal{B}$ induced by d. Moreover, we demand that $\mathcal{B}$ is compact, separable and connected. We denote the generic elements of $\mathcal{B}$ by x, y, z and open balls in $\mathcal{B}$ with the centre at x and the radius r as $B(x, r)$. The interpretation of metric d is as a measure of similarity of the alternative choices, i.e. $d(x, y) < d(x, z)$ means that x is more similar to y than to z (technically d measures the dissimilarity of the alternatives, but we call it a measure of similarity nevertheless).

In addition to that, we equip $\mathcal{B}$ with an arbitrary regular Borel measure λ, such that $\lambda(B(x, r)) > 0$ for all $x \in \mathcal{B}, r \in \mathbb{R}_{++}$. One simple and satisfactory example of such measure is to define $\lambda(B(x, r)) = r$. Abusing the notation a little, we also denote by λ the induced product measure on $\mathcal{B}^2$, which is given by $\lambda(A \times B) = \lambda(A)\lambda(B)$ and similarly by d the product metric on $\mathcal{B} \times \mathcal{B}$ given by $d((x_1, y_1), (x_2, y_2)) = \sqrt{d(x_1, x_2)^2 + d(y_1, y_2)^2}$.

We define Ω to be a set of all the permissible preference relations, meaning a set of all the binary relations on $\mathcal{B}$ that satisfy axioms 1–3 stated below. The generic element of Ω is denoted by ω.

Axiom 1. *(Rationality) Let $\omega \in \Omega$. Then ω is complete, reflexive and transitive.*

Axiom 2. *(Continuity) Let $\omega \in \Omega$. For each $x \in \mathcal{B}$ sets $\{y \in \mathcal{B} : x \succeq_\omega y\}$, $\{y \in \mathcal{B} : y \succeq_\omega x\}$ are closed in $\mathcal{B}$.*

Axiom 3. *(Limited Indifference) Let $\omega \in \Omega$ and $x \in \mathcal{B}$. Then*

$$\lambda(\{y \in \mathcal{B} : x \sim_\omega y\}) = 0.$$

The set Ω consists of all the relations that might be the preferences of a given consumer. Their unknown real preferences are the one element of Ω that is distinguished, and they are denoted by ω^*. We denote the relation of weak preference with respect to preferences $\omega \in \Omega$ by $x \succeq_\omega y$ and for strict preference and similarity relations we use respectively $\succ_\omega$ and $\sim_\omega$.

Axioms 1 and 2 are standard axioms of utility theory, and the theorem of Debreu (1954) states that preferences that satisfy those axioms can be represented by a continuous utility function. Axiom 3 states that indifference curves are "thin";

this is not necessary in any substantial way, and is it introduced only as a simplification. The intended consequence of this axiom is, that for a randomly chosen $x, y \in \mathcal{B}$ it should be very unlikely that $x \sim y$.

We use $\mathcal{D} \subset \mathcal{B}$ to denote a finite subset of alternative choices, that have already been consumed. By assumption, the real preference relations between all the pairs of elements of $\mathcal{D}$ are known. We denote by $\Omega(D) \subset \Omega$ those preference relations, that agree with ω^* on $\mathcal{D}$, that is $\Omega(\mathcal{D}) = \{\omega \in \Omega : \forall_{x,y \in \mathcal{D}} \quad x \succeq_\omega y \iff x \succeq_{\omega^*} y\}$.

To simplify the notation, we also define a set $\mathcal{K}_\mathcal{D}$ of formal expressions that correspond to those known relations, meaning the elements of $\mathcal{K}_\mathcal{D}$ are expressions of form xRy where $R \in \{\succ, \succeq, \sim, \prec, \preceq\}$ and x, y are elements of $\mathcal{D}$. The set $\mathcal{K}_\mathcal{D}$ consists of expressions that correspond to precisely those relations that agree with ω^*, so $x \succeq y \in \mathcal{K}_\mathcal{D} \iff x \succeq_{\omega^*} y$, and the same for other relation symbols. For the sake of brevity, we often present $\mathcal{K}_\mathcal{D}$ as for example $\mathcal{K}_\mathcal{D} = \{x_1 \succ \ldots x_i \sim x_{i+1} \succ \cdots \succ x_n\}$, but this definition has precisely the same meaning, as for example with the definition as above, we still have $x_n \preceq x_1 \in \mathcal{K}_\mathcal{D}$ (we never consider sets of relations that are inconsistent, meaning intransitive or irreflexive, as ω^* is transitive and reflexive).

We simplify the language a little by referring from now on to elements of $\mathcal{K}_\mathcal{D}$ as relations, or sometimes conditions, instead of "formal expressions that correspond to relations". For formal reasons we sometimes also consider sets of relations that are not linked to any particular $\mathcal{D}$, so for example we may consider $K = \{x_1 \succ x_2, x_3 \sim x_4\}$. With this definition, it still holds that for example $x_3 \succeq x_4 \in K$, meaning that K is a transitive closure of $\{x_1 \succ x_2, x_3 \sim x_4\}$. However we do not demand completeness, meaning that the definition does not specify any relation between x_1 and x_2 and it does not have to — just as x_1 does not have to be an element of $\mathcal{D}$ and we do not even have to specify $\mathcal{D}$ in this context. Moreover, for this K we also define $\Omega(K) = \{\omega \in \Omega : \forall_{xRy \in K} \quad xR_\omega y\}$.

Example 1. To visualise all these definitions and their intended interpretations, consider the example that the alternative choices are different movies that are uniquely and precisely described by some n-element vectors of characteristics, which are all normalised to $[0, 1]$, giving us $\mathcal{B} = [0, 1]^n$. Movies are similar if they have similar characteristics, which the consumer infers from the description and the trailer, and metric d, which we may assume is Euclidean, defines the specific shape of how the consumer perceives this similarity. The consumer has some real preferences ω^* over those movies, but does not know what those preferences are. However, they know how to rank the movies they have already seen, and vectors representing all those movies are the elements of $\mathcal{D}$, and their ranking of those movies is given by $\mathcal{K}_{\mathcal{D}}$. They also knows that their preferences have to be "sensible", where sensible is defined as satisfying axioms 1–3, and that they must agree with their ranking of the movies they have already seen, meaning they must be a member of $\Omega(\mathcal{D})$.

There are two important things to note from example 1. The first, is that ω^* is the only element that is interpreted as unknown to the consumer. Therefore we assume that the consumer knows all the alternative choices, remembers their past choices and ranks those choices without mistakes, though indeed their knowledge goes even deeper than that as they are also aware of axioms 1–3, meaning they are aware that they are, for example unlikely to find a movie that is exactly as good as Godzilla, or that if they prefer Godzilla to Rambo, then movies like Godzilla 2 that are sufficiently similar to Godzilla should also be preferred to Rambo. They also perceive the distances between alternatives, even those as yet unconsumed and do so without mistakes, so that if they perceive Godzilla 2 to be extremely similar to Godzilla then it cannot be the case that after watching Godzilla 2 they find the two not to be similar at all. It is possible that they find their tastes to be starkly different for Godzilla and Godzilla 2, but this would be the consequence of differences between the movies that they previously knew of but did not know their preferences for, rather than the other way around.

The second thing to note is, that metric d should be interpreted as the subjective perception of similarity by the consumer, and as such it should usually be assumed to be ex ante unknown in experimental studies. One consequence of this observation is the prospect of potential problems with empirical identification, as the probability measures that we introduce in section 3 depend on similarity between alternatives, so the question remains whether it is possible to identify jointly both the probability measure and the similarity metric behind it. Moreover, this perception of distance can be malleable when faced with framing or other marketing techniques (e.g. Mandel and Johnson 2002, Ariely et al. 2006).

The assumption that ω^* exists is mostly philosophical, and it has relatively few implications within the scope of this article. The only place during the whole article where we use ω^* is in the definition of the set of known relations $\mathcal{K}_{\mathcal{D}}$, and ω^* guarantees that the relations in $\mathcal{K}_{\mathcal{D}}$ are consistent with one another and stable. Keeping those two conditions constant, we could equally consider $\mathcal{K}_{\mathcal{D}}$ to be constructed from $\mathcal{D}$ via some procedure of preference construction and this would have no real impact on our analysis.

3 Construction of measures

This section is mostly technical and is oriented towards the development of probability measures on Ω, that could serve to represent the preferences that consumer with limited knowledge of their own tastes expects to have. To do this, we first need to introduce some topological structure on Ω. There are several studies that introduce topology on sets of preferences, among which Kannai (1970) is probably the best known. However, none of the constructions in the literature is sufficient for our purposes, as we need this topological structure to be linked to the sets that represent the information available to the consumer. We begin with the following definition.

Definition 1. *Let $x, y \in \mathcal{B}$ and $R \in \{\succ, \succeq, \sim, \prec, \preceq\}$. We define $[xRy] = \{\omega \in \Omega : xR_\omega y\}$ and call it a condition on x, y.*

For any finite sequence $(x_i, y_i, R_i)_{i=1}^n$ we denote the intersection of conditions $\bigcap_{i=1}^n [x_i R_i y_i]$ as $[\bigwedge_{i=1}^n x_i R_i y_i]$ and call it a conjunction of conditions on x_i, y_i and the union of conditions $\bigcup_{i=1}^n [x_i R_i y_i]$ as $[\bigvee_{i=1}^n x_i R_i y_i]$ and call it a disjunction of conditions on x_i, y_i. Similarly, for any $K = \{x_1 R_1 y_1, \ldots, x_n R_n y_n\}$ we denote $[K] = [\bigwedge_{i=1}^n x_i R_i y_i]$. The terms conjunction and disjunction are justified, as we have

$$\left[\bigwedge_{i=1}^n x_i R_i y_i\right] = \bigcap_{i=1}^n [x_i R_i y_i] = \left\{\omega \in \Omega : \bigwedge_{i=1}^n x_i R_{i\omega} y_i\right\},$$

and similarly

$$\left[\bigvee_{i=1}^n x_i R_i y_i\right] = \bigcup_{i=1}^n [x_i R_i y_i] = \left\{\omega \in \Omega : \bigvee_{i=1}^n x_i R_{i\omega} y_i\right\}.$$

For a given conjunction of conditions $[\bigwedge_{i=1}^n x_i R_i y_i]$, we also define its length by $l([\bigwedge_{i=1}^n x_i R_i y_i]) = n$ and a set $\mathrm{cp}([\bigwedge_{i=1}^n x_i R_i y_i]) = \{(x_i, y_i) : i \leq n\}$ of the pairs on which the conditions are imposed.

A little informally[1], let α be a logical formula constructed from expressions of the form $x_i R_i y_i$ for $R_i \in \{\succ, \succeq, \sim, \prec, \preceq\}$ together with logical connectives and quantifiers. As an example, we could define $\alpha = [(x_1 \succ x_2) \vee \neg(x_1 \sim x_3)] \implies [x_4 \preceq x_1]$. We denote the evaluation of α with respect to element $\omega \in \Omega$ by $\alpha_\omega \in \{0, 1\}$, that is $\alpha_\omega = 1$ if replacing every R_i in $x_i R_i y_i$ in α gives a true logical statement. For example, the statement $\alpha = [(x_1 \succeq x_2) \wedge (x_2 \succeq x_3)] \implies (x_1 \succeq x_3)$ evaluates to $\alpha_\omega = 1$ for every transitive ω (and would evaluate to $\alpha_\omega = 0$ for a non-transitive one) as $[(x_1 \succeq_\omega x_2) \wedge (x_2 \succeq_\omega x_3)] \implies (x_1 \succeq_\omega x_3)$ is a true statement for any transitive ω. For such logical formulas α we can similarly define $[\alpha] = \{\omega \in \Omega : \alpha_\omega = 1\}$.

We define the topology on Ω, to be the topology generated by the family of conditions $\{[x \succ y] : x, y \in \mathcal{B}\}$. As a consequence, closed sets in Ω are generated by the family of conditions $\{[x \succeq y] : x, y \in \mathcal{B}\}$.[2]

After those topological considerations, we are now ready to consider measures and corresponding σ-fields, on Ω, which we always denote by μ. The properties given by axioms 4-6 are the primitive assumptions that we make about the measure, and are kept constant through everything that follows.

Axiom 4. *(Non-degeneracy) Let $U \subset \Omega$ be open and nonempty. Then $\mu(U) > 0$.*

Axiom 5. *(Continuity) For all pairwise non-equal $x, y, x' \in \mathcal{B}$ and any $\epsilon > 0$ there exists $\delta > 0$ such that $d(x, x') < \delta \implies |\mu([x \succ y]) - \mu([x' \succ y])| < \epsilon$.*

Axiom 6. *(Restricted Indifference) Let $x, y \in \mathcal{B}$ be such that $x \neq y$. Then $\mu([x \sim y]) = 0$.*

Axiom 4 is very natural. We may interpret it as saying, that for any finite, consistent set of conditions, there is apriori a non-zero probability that those

[1]For a formal statement of this observation, we would need to introduce a whole machinery of first order logic and define a logical structure with Ω as a universe. We shy away from doing this in order to keep this section more readable, and we redirect anyone interested in how to do it to Enderton (2001) or any other introduction to mathematical logic.

[2]Even though it is abstract, the topology introduced in such a way is actually very natural. To see this, consider for each $\omega \in \Omega$ the function $f_\omega : \mathcal{B}^2 \to \{-1, 0, 1\}$, such that $f_\omega(x, y) = 1$ iff $x \succ y$, $f_\omega(x, y) = 0$ iff $x \sim y$ and $f_\omega(x, y) = -1$ iff $x \prec y$. For this representation to make sense, we equip the set $\{-1, 0, 1\}$ in the topology, such that $\emptyset, \{-1\}, \{1\}, \{-1, 1\}, \{-1, 0, 1\}$ are open, so with this definition $\omega \in \Omega$ is continuous if and only if f_ω is. Equipping the whole space of continuous functions from $\mathcal{B}^2$ into $\{-1, 0, 1\}$ in the standard product topology and embedding Ω into this space using the identification $\omega \to f_\omega$, now gives the same topology.

conditions are satisfied. Therefore, axiom 4 states, that for any pair $x, y \in \mathcal{B}$ with $x \neq y$ the consumer perceives that both $x \succ y$ and $y \succ x$ are apriori probable, though generally not with the same probability. The interpretation of axiom 5 is obvious, though it should be noted that continuity is restricted only to those elements of $\mathcal{B}$ that are not identical. To understand why, we should consider that $\mu([x \sim x]) = 1$ while axiom 6 means that for any $x' \neq x$ we have $\mu(x' \sim x) = 0$, so in this case we have an obvious discontinuity. For this reason, we abuse the notation a little and often treat $\mu([x \succ x]) = \frac{1}{2}$, making μ continuous everywhere in the sense of axiom 5. This simplification means we do not have to mention all the time that $x = y$ is a special case in every fact and definition we state.

Axiom 6 is again introduced to preserve simplicity by allowing us to ignore the possibility of indifference between alternatives, so we assume that the consumer does not consider indifference between the choices to be a legitimate possibility. It is clear that this axiom is related to axiom 3, but it is not necessarily implied by it. The connection is the other way around, as without axiom 3 it would be impossible to introduce this axiom on μ, as $[x \sim y]$ could have a non-empty interior. For the same reason, it would also be impossible to define the topology on Ω as we did without axiom 3.

Beside μ we also define $\mu_{\mathcal{D}}$, which is a conditional measure (given $\mathcal{D}$) on $\Omega(\mathcal{D})$ that corresponds to μ. Definition 2 specifies the connection between μ and $\mu_{\mathcal{D}}$.

Definition 2. *Let μ and some measurable $A \subset \Omega$ be given and denote $K = \mathcal{K}_{\mathcal{D}} \setminus \{x \sim y : x \sim y \in \mathcal{K}_{\mathcal{D}}\}$. We define*

$$\mu_{\mathcal{D}}(A) = \begin{cases} 0, & \textit{if } A \cap \Omega(\mathcal{D}) = \emptyset, \\ \frac{\mu(A \cap \Omega(K))}{\Omega(K)}, & \textit{otherwise.} \end{cases}$$

To understand definition 2, it should be noted that axiom 4 means it is never the case that $\mu(\Omega(\mathcal{D})) = 0$ if there is no such x, y that $x \sim y \in \mathcal{K}_{\mathcal{D}}$. This being the case, it simplifies to the standard definition of conditional probability. If there are some indifferences in $\mathcal{K}_{\mathcal{D}}$ however, definition 2 assigns values to sets as if those indifferences were not there. Those indifference relations are simply ignored when conditioning, but the support of $\mu_{\mathcal{D}}$ is still restricted to $\Omega(\mathcal{D})$, not $\Omega(K)$. It should be noted that by this definition, we have $x, y \in \mathcal{D} \implies \mu_{\mathcal{D}}([xRy]) \in \{0, 1\}$.

One important implication of this definition is that it explicitly assumes that the conditional measure, and by extension the preferences provided in section 4, is path or history independent, as the only thing that matters are the known relations provided by $\mathcal{K}_{\mathcal{D}}$, which are independent from the order in which the alternatives were explored. This is contrary to the assertions in the literature on preference construction (e.g. Payne et al. 1999) and is a consequence of our assumption that ω^* exists and is revealed by consumption.

We are mostly interested in the values of $\mu, \mu_{\mathcal{D}}$ when they are restricted to sets of conditions or their conjunctions and disjunctions. This is due to the natural interpretation of, for example, $\mu_{\mathcal{D}}([x \succ y])$ as the probability that x is better than y conditional on knowledge of $\mathcal{D}$; an alternative and equally correct interpretation would be as the probability that $\omega^* \in [x \succ y]$. Therefore we need to ensure that all the sets of conditions are actually measurable. Proposition 1 gives an obvious condition that is needed for this to happen.

Proposition 1. *Let σ_B denote a Borel σ-field on Ω and σ' be an arbitrary σ-field such that $\forall_{x,y \in \mathcal{B}} : [x \succ y] \in \sigma'$. Then $\sigma_B \subset \sigma'$.*

Proof. It suffices to note that $\sigma_B = \sigma(\{[x \succ y] : x, y \in \mathcal{B}\})$, which is obvious due to the definition of the topology on Ω.

□

Proposition 1 means we are from now on only interested in μ that are defined on a Borel σ-field and we assume it to be fixed and given; it should be noted that axioms 4-6 are well defined with such a σ-field. We do this even though proposition 1 allows us to take a strictly larger one, as the Borel σ-field is sufficient for our purposes.[3] Note that such a nice σ-field could not be obtained in proposition 1 without a topology closely linked to the information structure.

The following theorem 1 presents a tool that helps in defining measures. It essentially, states that to define a measure, it is only needed to assign in a coherent way the values for all the possible conjunctions of conditions. The statement of the theorem is preceded by some supporting facts and definitions.

Definition 3. *Fix $U \subset \Omega$ be given and let $K = \{\alpha_i : i \in I\}$ for arbitrary $I \subset \mathbb{N}$ and conjunctions of conditions α_i. We say that K is a representation of U if $U = \bigcup_{i \in I} \alpha_i$.*

[3] We never make use of the fact that our σ-field does not contain any non-Borel sets either, so the reader is free to choose their own favourite σ-field containing all the Borel sets.

Note that a representation of any set $U \subset \Omega$ as $U = [\bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}]$ is not unique, as defining $\alpha = \bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}$ clearly means $U = [(x \succ y) \vee (y \succeq x) \vee \alpha]$.

Definition 4. *We say that the representation $K = \{\alpha_i : i \in I\}$ is disjoint if $i_1 \neq i_2$ implies that $[\alpha_{i_1}] \cap [\alpha_{i_2}] = \emptyset$.*

Definition 5. *Let K_1, K_2 be two representations of some set $U \subset \Omega$. We say that K_1 is subordinate to K_2 if for all $\alpha \in K_1$ there is $\alpha' \in K_2$ such that $[\alpha] \subset [\alpha']$.*

Definitions 3–5 formalize basic concepts regarding disjunctions of conjunctions of conditions. Due to the construction of topology on Ω and definition 2 those are the sets that are of the most interest to us. The following lemma 1 shows that we can always represent any set U of this form using disjoint conditions.

Lemma 1. *Let $U = [\bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}]$ with $R_{ij} \in \{\succ, \succeq, \sim\}$. There exist such numbers m', n'_j and for all pairs $j \leq m', i \leq n_j$ some choice alternatives $(x'_{ij}, y'_{ij}) \in \mathcal{B}^2$ and relations $R'_{ij} \in \{\succ, \succeq, \sim\}$ such that representation $U = [\bigvee_{j=1}^{m'} \bigwedge_{i=1}^{n'_j} x'_{ij} R'_{ij} y'_{ij}]$ is disjoint.*

Proof. We proceed by induction on m, which is the number of conjunctions of conditions in the representation of U. Let $m = 1$. Then $U = [\bigwedge_{i=1}^{n_1} x_{i1} R_{i1} y_{i1}]$ and therefore the thesis is trivially satisfied. We just need to prove the implication that if the thesis is satisfied for some m, then it is satisfied for $m + 1$.

Assume, that for any $U = [\bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}]$ we have a disjoint representation as $U = [\bigvee_{j=1}^{m'} \bigwedge_{i=1}^{n'_j} x'_{ij} R'_{ij} y'_{ij}]$. Now assume $U = [\bigvee_{j=1}^{m+1} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}]$. By assumption, we have that $U = [\bigvee_{j=1}^{m'} \bigwedge_{i=1}^{n'_j} x'_{ij} R'_{ij} y'_{ij} \vee \bigwedge_{i=1}^{n_{m+1}} x_{im+1} R_{im+1} y_{im+1}]$. Denote $\gamma = \bigwedge_{i=1}^{n_{m+1}} x_{im+1} R_{im+1} y_{im+1} \wedge \bigwedge_{j=1}^{m'} \neg\alpha_j$, where $\neg\alpha_j = \bigvee_{i=1}^{n'_j} \neg x'_{ij} R'_{ij} y'_{ij}$ and $\neg x R y$ is, for R respectively $\succ, \succeq, \sim$, given by $y \succeq x$, $y \succ x$ and $(y \succ x \vee x \succ y)$, therefore γ is of the required form.

As every logical formula can be rewritten in disjunctive normal form, there exist some $(\alpha_k)_{k=1}^{m_k}$ such that $\alpha_k = \bigwedge_{i=1}^{n'_k} x'_{ik} R'_{ik} y'_{ik}$ and $\gamma = \bigvee_{k=1}^{m_k} \alpha_k$. To finish the proof, it suffices to show that each α_k is disjoint with each α_j, which follows trivially, as by construction we can show that for each k, j there exists some i and a logical formula ϕ such that we can represent $\alpha_k = \phi \wedge \neg x'_{ij} R'_{ij} y'_{ij}$.

$\square$

Definitions 6–8 define elementary operations on conjunctions of conditions and representations, that are necessary for the proof of the main result of this section.

Definition 6. *Let $\alpha = \bigwedge_{i=1}^{n} x_i R_i y_i$ be a logical formula and fix $x, y \in \mathcal{B}$. We call a pair of logical formulas $\alpha \wedge (x \succ y), \alpha \wedge (y \succ x)$ the partition of α by x, y. Moreover, let a finite disjoint representation K of some $U \subset \Omega$ be given, together with a finite set $A \subset \mathcal{B} \times \mathcal{B}$. We say $\tilde{K}$ is a full partition of K with respect to A if $\tilde{K}$ represents U and every $\alpha' \in \tilde{K}$ is of form $\alpha' = \alpha \wedge \left(\bigwedge_{(x,y)\in A} x R_{xy} y\right)$ where $\alpha \in K$ and $R_{xy} \in \{\succ, \prec\}$.*

Definition 7. *Let logical formulas γ_1, γ_2 of form $\alpha \wedge (x \succ y), \alpha \wedge (y \succ x)$ be given. We call logical formula α the merger of γ_1, γ_2. Moreover we say $K_1 = \{\alpha\}$ is a full merger of $K_2 = \{\alpha_1, \ldots, \alpha_n\}$ for $n > 1$ if K_2 is a disjoint representation of $U = [\alpha]$.*

Definition 8. *Let conjunction of conditions $\alpha = \bigwedge_{i=1}^{n} x_i \succ y_i$ and a finite set $A \subset \mathcal{B} \times \mathcal{B}$ be given. We define restriction of α to A as rest$(\alpha, A) = \bigwedge_{(x_i,y_i)\in A} x_i \succ y_i$.*

The following lemma 2 shows that any infinite representation of $U \subset \Omega$ can be obtained from some finite representation using a sequence of partitions and mergers and lemma 3 shows that for two different representations of U, a very natural condition on the values assigned by some set function μ_0 suffices so that the value of $\mu_0(U)$ is independent from the choice of the representation of U.

Lemma 2. *Let $U \subset \Omega$ and fix two disjoint representations $K_0 = \{\alpha_1, \ldots, \alpha_n\}$ and $K = \{\alpha'_j : j \in \mathbb{N}_+\}$ of U such that K is subordinate to K_0. There exists a sequence $(K_l)_{l\in\mathbb{N}}$ of representations such that $\bigcap_{k=0}^{\infty} \bigcup_{l=k}^{\infty} K_l = K$ and that K_{l+1} is obtained from K_l using only partitions and mergers (full or otherwise).*

Proof. Fix U, K_0 and K as in the statement of the lemma, and assume all elements of K_0 are enumerated, meaning that $K_0 = \{\alpha_1, \ldots, \alpha_n\}$. We prove this lemma constructively, by providing a procedure to obtain a sequence of representations that satisfy the conditions of the lemma. For each $l \in \mathbb{N}$ we do the following steps. Fix $\alpha_i \in K_l$, with $i = 1$ in case $l = 1$. We also fix j, starting with $j = 1$ for $l = 1$. case $l = 1$ Define $K_{\alpha_i} = \{\alpha' \in K : [\alpha'] \subset [\alpha_i]\}$ and $K_{\alpha_i}(n) = \{\alpha' \in K_{\alpha_i} : l(\alpha') = n\}$. Fix n to be a smallest such number that $K_{\alpha_i}(n) \neq \emptyset$ and denote

$$D_{\alpha_i}(n) = \bigcup_{\alpha' \in K_{\alpha_i}(n)} \mathrm{cp}(\alpha') \setminus \mathrm{cp}(\alpha_i),$$

so that $D_{\alpha_i}(n)$ is a set of points on which additional conditions in $K_{\alpha_i}(n)$ are imposed. Note, that $D_{\alpha_i}(n)$ is finite. Define $K_{\alpha_i}^{FP}$ to be a full partition of α with respect to $D_{\alpha_i}(n)$ and define $K_{\alpha_i}^{rest} = \{\text{rest}(\alpha', D_{\alpha_i}(n) \cup cp(\alpha_i)) : \alpha' \in K_{\alpha_i}\}$.

As both $K_{\alpha_i}^{FP}, K_{\alpha_i}^{rest}$ are finite and disjoint and $K_{\alpha_i}^{FP}$ is subordinate to $K_{\alpha_i}^{rest}$, we can obtain each element $\alpha' \in K_{\alpha_i}^{rest}$ by a full merger of all elements of $K_{\alpha_i}^{FP}$ that satisfy $[\alpha''] \subset [\alpha]$. As a result, we can obtain $K_{\alpha_i}^{rest}$ from α using only partitions and mergers.

Finally, define $K_{l+1} = K_l \cup K_{\alpha_i}^{rest} \setminus \{\alpha_i\}$. We do not change the enumeration of elements in K_{l+1}. As such, all elements of $K_{l+1} \cap K_{\alpha_i}^{rest}$ are not enumerated (for now), meaning that there is no i' such that $\alpha_{i'} \in K_{l+1} \cap K_{\alpha_i}^{rest}$. Note that by construction $K_{\alpha_i}(n) \subset K_{l+1}$, K_{l+1} is disjoint and K is subordinate to K_{l+1}.

Now, if $\alpha_{i+1} \in K_{l+1}$, increase i, l by one and perform the same operations as we did up to this point. In the other case, increase l and j by one, set $i = 1$ and enumerate all elements of K_l.

Note, that clearly if $\alpha' \in K$ and $\alpha' \in K_l$ for any l then also $\alpha' \in K_{l+1}$. Therefore in order to finish the proof, we just need show that every element $\alpha' \in K$ is obtained as an element of K_l for some l. Let $[\alpha'] \subset [\alpha]$ for some $\alpha \in K_0$ and fix $n_{\alpha'} = |\{n \le l(\alpha') : K_\alpha(n) \neq \emptyset\}|$.

We claim, that we obtain α' as an element of K_l for some l such that $j = n_{\alpha'}$. Indeed, if $n_{\alpha'} = 1$ we have already shown it. Consider $n_{\alpha'} > 1$. There is i such that $[\alpha'] \subset [\alpha_i]$ for $j = 1$ and there is $\alpha'' \in K_{\alpha_i}^{rest}$ such that $[\alpha'] \subset [\alpha'']$. As $K_{\alpha''} \subset K_{\alpha_i} \setminus K_{\alpha_i}(n)$ where n is the smallest number such that $K_{\alpha_i}(n) \neq \emptyset$ we get that $|\{n \le l(\alpha') : K_{\alpha''}(n) \neq \emptyset\}| < n_{\alpha'}$, proving our claim. As for each j we perform a finite number of partitions and mergers, the proof is finished.

□

Lemma 3. *Fix two finite disjoint representations* $K_1 = \{\alpha_1, \ldots, \alpha_{m_1}\}$ *and* $K_2 = \{\alpha'_1, \ldots, \alpha'_{m_2}\}$ *of some set* $A \subset \Omega$. *Assume that set function*

$$\mu_0 : \{A \subset \Omega : \exists_{x_i, y_i, n}\ A = [\bigwedge_{i=1}^{n} x_i \succ y_i]\} \to [0, 1]$$

that for any conjunction of conditions α *satisfies* $\mu_0([\alpha \wedge (x \succ y)]) + \mu_0([\alpha \wedge (y \succ x)]) = \mu_0([\alpha])$ *is given. Then* $\sum_{j=1}^{m_1} \mu_0([\alpha_j]) = \sum_{j=1}^{m_2} \mu_0([\alpha'_j])$.

Proof. By the condition that $\mu_0([\alpha \wedge x \succ y]) + \mu_0([\alpha \wedge y \succ x]) = \mu_0([\alpha])$ the values of μ_0 are assigned in such a way that replacing any α_{j_0} or α'_{j_0} by its arbitrary partition, for example replacing α_{j_0} by $\alpha^1_{j_0}, \alpha^2_{j_0}$ gives $\sum_{j=1}^{m_1} \mu_0([\alpha_j]) = \mu_0([\alpha^1_{j_0}]) + \mu_0([\alpha^2_{j_0}]) + \sum_{j \neq j_0}^{m_1} \mu_0([\alpha_j])$. Therefore it suffices to show, that there exists a finite sequence of partitions from both K_1 and K_2 to some $K = \{\alpha^{l_1}_1, \ldots, \alpha^{l_1}_{k_1}\}$, meaning we can obtain the same finite subset of formulas as a result of the recursive partitioning of K_1 and K_2. It suffices to define

$$D = \bigcup_{j=1}^{m_1} \mathrm{cp}(\alpha_j) \cup \bigcup_{j'=1}^{m_2} \mathrm{cp}(\alpha'_{j'}),$$

and fix K to be a representation obtained by partitioning of all formulas in K_1 on all elements of D. Obviously, partitioning all elements of K_2 on all elements of D we also obtain K and the proof is finished.

$\square$

Now we are ready to prove the main theorem of this section.

Theorem 1. *Assume, that for all $n \in \mathbb{N}_+$ and all sequences $(x_i, y_i)_{i=1}^n \in \mathcal{B} \times \mathcal{B}$ the values of set function $\mu_0([\bigwedge_{i=1}^n x_i \succ y_i]) > 0$ are given and satisfy*

$$\mu_0([\bigwedge_{i=1}^{n-1} x_i \succ y_i \wedge x_i \succ y_i]) + \mu_0([\bigwedge_{i=1}^{n-1} x_i \succ y_i \wedge y_i \succ x_i]) = \mu_0([\bigwedge_{i=1}^{n-1} x_i \succ y_i),$$

and

$$\mu_0([x \succ y]) + \mu_0([y \succ x]) = 1.$$

There then exists a unique probabilistic measure μ defined on the whole Borel σ-field of Ω such that for all conjunctions of conditions we have $\mu_0([\bigwedge_{i=1}^n x_i \succ y_i]) = \mu([\bigwedge_{i=1}^n x_i \succ y_i])$.

Proof. Define a family of sets

$$\mathcal{A} = \left\{ [\bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}] : x_{ij}, y_{ij} \in \mathcal{B}, R_{ij} \in \{\succ, \succeq, \sim\} \right\} \cup \emptyset.$$

We show, that $\mathcal{A}$ is an algebra of sets. It contains an empty set, and is obviously closed under binary unions. Moreover, it is closed under complementation, as

$$[\bigvee_{j=1}^{m} \bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij}] \setminus [\bigvee_{j=1}^{m'} \bigwedge_{i=1}^{n'_j} x'_{ij} R'_{ij} y'_{ij}] = \left[\bigvee_{j=1}^{m} \left(\bigwedge_{i=1}^{n_j} x_{ij} R_{ij} y_{ij} \wedge \bigwedge_{j=1}^{m'} \bigvee_{i=1}^{n'_j} \left(\neg x'_{ij} R'_{ij} y'_{ij} \right) \right) \right] = A,$$

where $\neg x'_{ij} \succ y'_{ij} = y'_{ij} \succeq x'_{ij}$, $\neg x'_{ij} \succeq y'_{ij} = y'_{ij} \succ x'_{ij}$ and $\neg x'_{ij} \sim y'_{ij} = y'_{ij} \succ x'_{ij} \vee x'_{ij} \succ y'_{ij}$. As every logical formula can be stated in disjunctive normal form, $A \in \mathcal{A}$ and we find that $\mathcal{A}$ is an algebra of sets.

We first extend μ_0 to the whole $\mathcal{A}$ as follows: define $\mu_0([\bigwedge_{i=1}^n x_i R_i y_i]) = 0$ if any $R_i = \sim$ and $[\bigwedge_{i=1}^n x_i R_i y_i] = [\bigwedge_{i=1}^n x_i \succ y_i]$ otherwise. Moreover from proposition 1 we get that each $A \in \mathcal{A}$ can be represented some $[\bigvee_{j=1}^m \bigwedge_{i=1}^{n_j} x_{ij} \succ y_{ij}]$ that is disjoint. Therefore we define $\mu_0(A) = \sum_{j=1}^m \mu_0([\bigwedge_{i=1}^{n_j} x_{ij} \succ y_{ij}])$. Note that this is well defined due to lemma 3, which we can apply due to the condition in the statement of the theorem. This is therefore a unique extension of μ_0 to $\mathcal{A}$ such that the extended μ_0 is finitely additive.

We need to show, that μ_0 is a pre-measure on $\mathcal{A}$. Fix some $A \in \mathcal{A}$ and let $(A_j)_{j=1}^\infty$, $A_j \in \mathcal{A}$ be disjoint and such that $\bigcup_{j=1}^\infty A_j = A$. From proposition 1 and finite additivity of μ_0 we can assume without loss of generality that A_j are of form $\alpha_j = [\bigwedge_{i=1}^{n_j} x_{ij} \succ y_{ij}]$. Also, denote by K the representation of A corresponding to A_j's, so that $K = \{\alpha_j : j \in \mathbb{N}_+\}$. We need to show that $\mu_0(A) = \sum_{j=1}^\infty \mu_0(A_j)$.

Let $\tilde{K_0}$ be an arbitrary disjoint representation of A. By lemma 1 some disjoint representation exists. Define $D = \bigcup_{\alpha \in \tilde{K_0}} \mathrm{cp}(\alpha)$ and take $K_0 = \{\mathrm{rest}(\alpha) : \alpha \in K\}$. Clearly, K_0 is finite and K is subordinate to K_0. Therefore from lemma 2 we have that there exist a sequence of representations $(K^l)_{l \in \mathbb{N}}$ such that K^{l+1} is obtained from K_l using mergers and partitions only, and that $\bigcap_{k=0}^\infty \bigcup_{l=k}^\infty K^l = K$, so the limit of this sequence of recursive partitions and mergers is K. Note, that by finite additivity of μ_0 mergers and partitions have no impact, meaning that for every $l \in \mathbb{N}$

$$\sum_{\alpha \in K^l} \mu_0([\alpha]) = \mu_0(A).$$

Now consider two sequences $m_l = \sum_{\alpha \in K^l} \mu_0([\alpha])$ and $m^l = \mu_0([\bigcap_{k=0}^l \bigcup_{k'=k}^l K^{k'}])$. It is clear that both m_l and m^l are constant and equal to $\mu_0(A)$. Therefore $\lim_{l\to\infty} m_l = \mu_0(A) = \lim_{l\to\infty} m^l$. It now suffices to note, that $\lim_{l\to\infty} m_l = \sum_{\alpha \in K} \mu_0([\alpha])$ and $\lim_{l\to\infty} m^l = \lim_{l\to\infty} \mu_0([\bigcap_{k=0}^l \bigcup_{k'=k}^l K^{k'}]) = \mu_0(\bigcup_{\alpha \in K}[\alpha]) = \mu_0(A)$. Therefore μ_0 is a pre-measure on $\mathcal{A}$.

As $\mathcal{A}$ is an algebra of sets and μ_0 is a finite pre-measure that is uniquely extended to $\mathcal{A}$ from given values, then by Caratheodory's extension theorem (e.g. Ash 1999), it follows, that there exists a unique σ-finite measure μ that extends μ_0 to the whole σ-field generated by $\mathcal{A}$. As $\mathcal{A}$ contains the generating set of the topology on Ω, the σ-field generated by it must contain all open sets, and as a consequence all Borel sets.

Now to finish the proof, it suffices to show that μ is probabilistic, but this follows trivially from the condition that $\mu_0([x \succ y]) + \mu_0([y \succ x]) = 1$.

$\square$

Theorem 1 is very important. It shows that we can define measures μ on the whole Ω by their values on sets defined by strict conjunctions of conditions only. As those are precisely the sets that are of interest to us, this is a very nice property to have. Notably, the values of the measure on those sets also determine the indirect learning of the consumer. In order to see it, consider that indirect learning essentially describes the correlation between relations, meaning that if from learning that $x_1 \succ y_1$ we indirectly learn that probably $x_2 \succ y_2$, it is due to the fact, that $x_1 \succ y_1$ and $x_2 \succ y_2$ are correlated, that is the majority (with respect to μ) of preferences in $\Omega([x_1 \succ y_1])$ also satisfy $x_2 \succ y_2$, whereas conversely the majority in $\Omega([y_1 \succ x_1])$ would satisfy $y_2 \succ x_2$. Specification of $\mu_0([x_1 \succ y_1 \wedge x_2 \succ y_2])$ with $\mu_0([x_1 \succ y_1])$ and $\mu_0([x_2 \succ y_2])$ already specified means doing precisely that – specifying the strength of the correlation between these two relations. As a consequence, theorem 1 shows that there are no bounds whatsoever on indirect learning in our model.

Theorem 1 also allows us to define the measures using conditional probabilities that we can specify independently, as shown by corollary 1.

Corollary 1. *Let for all $x, y \in \mathcal{B}$ and for all sets of conditions K values of $\mu^0_K([x \succ y])$ be given and satisfy*

$$\mu^0_K([x \succ y]) + \mu^0_K([y \succ x]) = 1, \quad \mu^0_\emptyset([x \succ y]) + \mu^0_\emptyset([y \succ x]) = 1.$$

There exists a unique probabilistic measure μ defined on the whole Borel σ-field, such that for all x, y we have $\mu^0_\emptyset([x \succ y]) = \mu([x \succ y])$ and $\mu^0_K([x \succ y]) = \mu_K([x \succ y])$.

Proof. Note that $\mu([\bigwedge_{\alpha \in K} \alpha \wedge x \succ y]) = \mu_K([x \succ y])\mu(\Omega(K))$. Now for it to follow from theorem 1 we just need to show that we are able to calculate $\mu(\Omega(K))$ using values of $\mu_K([x \succ y])$ only. It suffices to do this inductively. Let $K = \{x_1 \succ x_2\}$. Then $\mu(\Omega(K)) = \mu_\emptyset([x_1 \succ x_2])$. Now assume we are given $\mu(\Omega(K))$ for $K = \{x_1 \succ \cdots \succ x_n\}$ and let $K' = \{x_1 \succ \cdots \succ x_{n+1} \succ \cdots \succ x_n\}$. From definition 2 we have $\mu(\Omega(K')) = \mu(\Omega(K))\mu_K([x_j \succ x_{n+1} \wedge x_{n+1} \succ x_{j+1}])$, where $\mu_K([x_j \succ x_{n+1} \wedge x_{n+1} \succ x_{j+1}]) = 1 - \mu_K([x_{n+1} \succ x_j \vee x_{j+1} \succ x_{n+1}])$. Given that $x_j \succ x_{j+1} \in K$, the sets $[x_{n+1} \succ x_j], [x_{j+1} \succ x_{n+1}]$ are disjoint in $\Omega(K)$. Therefore

$\mu_K([x_j \succ x_{n+1} \wedge x_{n+1} \succ x_{j+1}]) = 1 - \mu_K([x_{n+1} \succ x_j]) - \mu_K([x_{j+1} \succ x_{n+1}])$ and the proof is finished.

$\square$

Note, that the layers given by the sets of form $[x \succ y]$ on which we define $\mu^0_{\mathcal{D}}$ in corollary 1 are independent of each other for different $\mathcal{D}$. Indeed, the only condition is, that $\mu^0_{\mathcal{D}}([x \succ y]) + \mu^0_{\mathcal{D}}([y \succ x]) = 1$. Therefore we can sometimes restrict our attention to some fixed $\mathcal{D}$ and consider $\mu_{\mathcal{D}}$ as a probability distribution on $\mathcal{B}^2$, ignoring what happens otherwise. Corollary 6 says precisely that we can then "glue" those distributions into a Borel measure on Ω.

Even though neither of theorem 1 or corollary 1 actually provide us with a proof that some probabilistic measure μ on Ω satisfying axioms 4-6 actually exists (this is finally proven in section 4 by theorem 2), these results give us the tools for defining such measures. In a static case we do not use the whole measure μ much, with conditional measures for some given $\mathcal{D}$ proving sufficient for our purposes. However, theorem 1 is of paramount importance, as it shows that we can think of consumer choices given all possible $\mathcal{D}$ as being generated by a common decision process that is informed by a common probability measure that glues together all the conditional measures.

The conditions given by theorem 1 for which assignments of conditional measures we can glue together with different possible $\mathcal{D}$ are very weak. As an upside, this allows us through the following sections to concentrate on the constructions of $\mu_{\mathcal{D}}$ with minimal regard as to how they correspond to one another. A downside is that certainly not all assignments of $\mu_{\mathcal{D}}$ are reasonable as representations of changes in the perception by the consumer of their own preferences. Any further work on this topic that extends beyond the static case we currently consider should consider imposing additional restrictions on μ in that regard.

4 Conditional preferences

In this section, we turn towards the question of how a consumer equipped with a Borel probability measure μ on Ω perceives their preferences. We focus on the static setting, so we assume $\mathcal{D}$ is fixed throughout. As noted before, we also leave aside choice functions, and explore instead different decision modes. Moreover, theorem 1 lets us mostly ignore μ and focus on $\mu_{\mathcal{D}}$ only.

As a basic distinction, we assume that the consumer has two conflicting motivations that we call the consumptional and the experimental: consumptional motivation comes from the immediate satisfaction from the consumer's choice, and experimental motivation comes from the consumer getting to know their real preferences better. For now, we consider purely the consumptional motivation. It seems natural that the consumer who is preoccupied with their immediate satisfaction from consumption should try to guess their real preferences ω^*. The most obvious way to do this is given by definition 9.

Definition 9. *Let $\mu_{\mathcal{D}}$ be given. We call relation $\succeq_{\mathcal{D}}$ (also denoted $\omega_{\mathcal{D}}$)) defined by $x \succeq_{\mathcal{D}} y \iff \mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2}$ a conditional preference relation, which is conditional on $\mathcal{D}$. Accordingly, for a given preference relation $\succeq_{\mathcal{D}}$ we say that $\mu_{\mathcal{D}}$ such that $x \succeq_{\mathcal{D}} y \iff \mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2}$ represents $\succeq_{\mathcal{D}}$ as conditional preferences, that are conditional on $\mathcal{D}$.*

When $\mathcal{D}$ is fixed we simply say that the conditional preferences of the consumer are given by $\omega_{\mathcal{D}}$, and when the context is clear we simply say that $\mu_{\mathcal{D}}$ represents $\omega_{\mathcal{D}}$. Even though we consider other decision modes, conditional preferences are the primary one, and the other modes originate from the study of the properties of the conditional preferences. We perceive $\omega_{\mathcal{D}}$ to be the most similar approach to the substantively rational approach to consumer choice. Note that for the time being we do not know whether $\omega_{\mathcal{D}}$ is actually a preference relation. The following definitions, 10 and 11, are of primary importance for the study of conditional preferences.

Definition 10. *We say that $\mu_{\mathcal{D}}$ is coherent if $(\mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2}) \implies (\forall_{z \in \mathcal{B}} : \mu_{\mathcal{D}}([x \succeq z]) \geq \mu_{\mathcal{D}}([y \succeq z]))$.*

Definition 11. *We say that $\mu_{\mathcal{D}}$ is weakly coherent if $(\mu_{\mathcal{D}}([x \succeq y]) = \frac{1}{2}) \implies (\forall_{z \in \mathcal{B}} : \mu_{\mathcal{D}}([x \succeq z]) \geq \frac{1}{2} \iff \mu_{\mathcal{D}}([y \succeq z]) \geq \frac{1}{2})$.*

The coherence of $\mu_{\mathcal{D}}$ is a property that would be nice to have, as it means that the notion of how likely it is that x is better than y has some significance. For incoherent measures, the only thing that matters is whether this probability is greater or smaller than half. As a consequence we can compare alternative choices indirectly for coherent measures, using any $z \in \mathcal{B}$ as a reference point. This observation is formally stated by proposition 2.

Proposition 2. *Let $z \in \mathcal{B}$ and $\succeq_z$ (also denoted ω_z) be defined as $x \succeq_z y \iff \mu_{\mathcal{D}}([x \succeq z]) \geq \mu_{\mathcal{D}}([y \succeq z])$. The following are all true.*

1. *Let $z \in \mathcal{B}$. Then ω_z is complete, reflexive, transitive and continuous.*

2. *Let $z \in \mathcal{B}$. Then $\forall_{z' \in \mathcal{B}} : \omega_z = \omega_{z'}$ if and only if $\mu_{\mathcal{D}}$ is coherent.*

3. *We have that $\forall_{z \in \mathcal{B}} : \omega_z = \omega_{\mathcal{D}}$ if and only if $\mu_{\mathcal{D}}$ is coherent.*

Proof. The first point of the proposition is obvious, as $u_z(x) = \mu_{\mathcal{D}}([x \succeq z])$ is a continuous function[4] that clearly represents ω_z and Debreu's (1954) theorem states that such a utility function exists if and only if ω_z is complete, reflexive, transitive and continuous.

We now prove point 2. Assume $\mu_{\mathcal{D}}$ is coherent and fix arbitrary $x, y \in \mathcal{B}$ such that $\mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2}$. Coherence means that for any $z_1, z_2 \in \mathcal{B}$ we have that $u_{z_1}(x) \geq u_{z_1}(y)$ and $u_{z_2}(x) \geq u_{z_2}(y)$, and therefore $\omega_{z_1} = \omega_{z_2}$. Now assume that $\forall_{z_1, z_2 \in \mathcal{B}} : \omega_{z_1} = \omega_{z_2}$, and we show that $\mu_{\mathcal{D}}$ is coherent. Again fix $x, y \in \mathcal{B}$ such that $\mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2}$. Note that for any $z \in \mathcal{B}$, u_z and u_y represent the same preferences. As we assume that $u_y(x) \geq u_y(y)$, we therefore get that $u_z(x) \geq u_z(y)$, and therefore coherence.

Finally, we prove 3. Assume that $\forall_{z \in \mathcal{B}} : \omega_z = \omega_{\mathcal{D}}$. Therefore especially for any $z_1, z_2 \in \mathcal{B}$ we have $\omega_{z_1} = \omega_{z_2}$, therefore following point 2 $\mu_{\mathcal{D}}$ is coherent. Now assume $\mu_{\mathcal{D}}$ is coherent, so from point 2 for all $z_1, z_2 \in \mathcal{B}$ we have that $\omega_{z_1} = \omega_{z_2}$. Therefore especially for any z we have $\omega_y = \omega_z$ and $\mu_{\mathcal{D}}([x \succeq y]) \geq \frac{1}{2} \implies u_y(x) \geq u_y(y)$, and therefore $\forall_{z \in \mathcal{B}} : \omega_z = \omega_{\mathcal{D}}$.

□

The relation ω_z defined by proposition 2, which we call the indirect preference relation (with respect to $\mathcal{D}$ and z, omitting $\mathcal{D}$ if it is clear from context) is the second decision mode to be introduced to our study. Following proposition 2, the indirect choice of an alternative using $z \in \mathcal{B}$ as a reference point is both independent from z and equivalent to the choice using direct comparisons given by $\omega_{\mathcal{D}}$ if and only if $\mu_{\mathcal{D}}$ is coherent. This is an important observation, as it is easy to give an example of a situation, where it could be considered reasonable to use indirect choice.

[4] As noted in section 3 function u_z is continuous due to our definition that $\mu_{\mathcal{D}}([x \succ x]) = \frac{1}{2}$.

Example 2. One especially fitting example concerns a consumer who is very averse to uncertainty. Consider a foreign traveller, who comes to a shoddy-looking roadside restaurant. Every dish on the menu is new to them, but they can understand the descriptions. They are hungry, and just want to be sure that whatever they order, is going to be edible. Edible is defined as anything that is at least as good as some known dish x. The consumer's choice function is therefore, to maximise the probability of their choice being as good as x. Formally, denoting the menu in the restaurant as A, this function is $\operatorname{argmax}_{y \in A}\{\mu_{\mathcal{D}}([y \succeq x])\}$. The solution to the choice problem, is therefore clearly given by the maximisation of u_x on A. Note that if the consumer's $\mu_{\mathcal{D}}$ is incoherent, it is possible that their choice will be different from the one that they would make if they used $\omega_{\mathcal{D}}$, as they might choose to eat fish and chips rather than beef guts even though they suspect beef guts will be actually better then fish and chips, just because there is also a higher probability that the beef guts will be completely inedible. Moreover, note that if $\mu_{\mathcal{D}}$ is incoherent, the choice is reference point dependent, so it is entirely possible, that if the consumer used some other x', even such that $x' \sim_{\mathcal{D}} x$, they would consider beef guts to be actually less risky than fish and chips, and change their choice accordingly, or be unable to decide if they actually tried to consider both reference points.

The utility functions u_z representing ω_z that we used in the proof of proposition 2 are a very useful tool to work with, so we define them formally in definition 12.

Definition 12. *Let $z \in \mathcal{B}$ be given. We denote $u_z(x) = \mu_{\mathcal{D}}([x \succ z])$ and call this u_z an indirect utility function with respect to z, and we call z its reference point.*

As example 2 shows, coherence would be a very nice property to have, since with a coherent measure it would essentially be pretty hard to think of an example where an intuitively rational consumer chooses differently to how $\omega_{\mathcal{D}}$ dictates. However it is not to be. As shown by proposition 3, $\mu_{\mathcal{D}}$ are almost never coherent.

Proposition 3. *Let $x_1, x_2, x_3 \in \mathcal{D}$ such that $x_1 \succ x_2, x_2 \succ x_3 \in \mathcal{K}_{\mathcal{D}}$. Then $\mu_{\mathcal{D}}$ is not coherent.*

Proof. Fix $\mathcal{D}$ as in the statement of the theorem. As $\mu_{\mathcal{D}}([x_1 \succeq x_3]) = 1$ and $\mu_{\mathcal{D}}([x_2 \succeq x_3]) = 1$ from continuity for any disjoint open ball $B(x_1, r_1), B(x_2, r_2) \subset \mathcal{B}$ and for any $z_2 \in B(x_2, r_2)$ there exists $z_1 \in B(x_1, r_1)$ such that $\mu_{\mathcal{D}}([z_2 \succeq$

$x_3]) > \mu_{\mathcal{D}}([z_1 \succeq x_3])$. Therefore from coherence $\mu_{\mathcal{D}}([z_2 \succeq z_1]) \geq \frac{1}{2}$. However $\mu_{\mathcal{D}}([x_1 \succeq x_2]) = 1$ and therefore from continuity there exist disjoint open balls $B(x_1, r_1), B(x_2, r_2) \subset \mathcal{B}$ so that for all $z_1 \in B(x_1, r_1), z_2 \in B(x_2, r_2)$ we have $\mu_{\mathcal{D}}([z_1 \succ z_2]) > \frac{1}{2}$, which is a contradiction.

□

Note that beside showing that $\mu_{\mathcal{D}}$ can only possibly be coherent in the very special case of $|\{x \succ y : x, y \in \mathcal{D}\}| \leq 1$, the proof of proposition 3 hints towards the source of this lack of coherence. Incoherence is introduced to $\mu_{\mathcal{D}}$ by experience, meaning that when $x_1 \succ x_2 \in \mathcal{K}_{\mathcal{D}}$ then $\mu_{\mathcal{D}}([x_1 \succ x_2]) = 1$, combined with the continuity of $\mu_{\mathcal{D}}$. Combined with proposition 2 we find as an additional consequence of proposition 3 that for $|\mathcal{D}| > 2$ there always exists such $z_1, z_2 \in \mathcal{B}$ such that $\omega_{\mathcal{D}} \neq \omega_{z_1}$ and $\omega_{z_1} \neq \omega_{z_2}$.

Proposition 2 informs us that for a coherent $\mu_{\mathcal{D}}$, relation $\omega_{\mathcal{D}}$ is a preference relation, meaning it is complete, transitive, reflexive and continuous. However, it is not a necessary condition, as shown by proposition 4.

Proposition 4. *Conditional preferences $\omega_{\mathcal{D}}$ are transitive if and only if $\mu_{\mathcal{D}}$ is weakly coherent.*

Proof. Note that $\omega_{\mathcal{D}}$ is transitive if and only if its lower contour sets $LCS_x = \{y \in \mathcal{B} : x \succeq_{\mathcal{D}} y\}$ are such that $x_1 \succeq_{\mathcal{D}} x_2 \iff LCS_{x_2} \subset LCS_{x_1}$. As weak coherence is a strictly weaker condition, a transitive $\omega_{\mathcal{D}}$ clearly implies weak coherence. Therefore we only need to prove that the weak coherence of $\mu_{\mathcal{D}}$ implies the transitivity of $\omega_{\mathcal{D}}$.

Fix some $x, y \in \mathcal{B}$ such that $x \succ_{\mathcal{D}} y$ and assume there exists such $x' \prec_{\mathcal{D}} y$ such that $x' \succeq_{\mathcal{D}} x$. Without a loss of generality, we can assume that $x' \sim_{\mathcal{D}} x$, as if $x' \succ_{\mathcal{D}} x$ and $y \succ_{\mathcal{D}} x'$ we have from the continuity of $\mu_{\mathcal{D}}$ that there must exist some x'' on the path connecting x' and y such that $x'' \sim_{\mathcal{D}} x$ and $x'' \prec y$ (such a path exists, as $\mathcal{B}$ is connected by assumption). Now we have that $\mu_{\mathcal{D}}(x \succeq x') = \frac{1}{2}$ and $\mu_{\mathcal{D}}(x' \succeq y) < \frac{1}{2}$, therefore by weak coherence it must be that $\mu_{\mathcal{D}}(x \succeq y) < \frac{1}{2}$, which contradicts the assumption that $x \succ_{\mathcal{D}} y$.

□

From proposition 4 we know that the transitivity of $\omega_{\mathcal{D}}$ is equivalent to weak coherence. To understand better what weak coherence actually means, we contrast it with normal, meaning not weak, coherence. It may be recalled that coherence essentially means, that for all the reference points $z \in \mathcal{B}$ the rankings of alternative choices given by indirect comparisons with z are the same. Therefore if $u_{z_1}(x) > u_{z_1}(y)$ it must be that $u_{z_2}(x) > u_{z_2}(y)$. With weak coherence, this is no longer the case. The prime reason for this is that weak coherence is only binding for the case of indifferent goods, so for it to say anything at all about the rankings above, it would have to be that $z_1 \sim_{\mathcal{D}} z_2$. Moreover, even for $z_1 \sim_{\mathcal{D}} z_2$ it actually allows for the case where $u_{z_1}(x) > u_{z_1}(y)$ and $u_{z_2}(x) < u_{z_2}(y)$. The only situation it prohibits is when some alternative choices switch sides with relation to the indifference curve, so it is impossible to have $u_{z_1}(x) > \frac{1}{2}$ and $u_{z_2}(x) \leq \frac{1}{2}$. Note that this differs from previous considerations, as in the prohibited situation we only consider one alternative choice x instead of both alternatives x, y. This is very much the essence of the difference between coherence and weak coherence. Choosing some reference points z_1, z_2, coherence demands that the relations between any pair of alternatives x, y is the same when reference points are changed. However, weak coherence is only interested in the relations between one alternative and the reference point itself.

From the viewpoint of indirect comparisons, the transitivity of $\omega_{\mathcal{D}}$ for a weakly coherent $\mu_{\mathcal{D}}$ is quite peculiar. To illustrate this point further, consider the following, example 3.

Example 3. Assume a consumer is interested in choosing a movie to watch, and narrows the list down to three titles only: Godzilla, Titanic and Gone with the Wind, denoted respectively as G, T and W. For a fully coherent $\mu_{\mathcal{D}}$, the indirect rankings must be equal, so assume therefore that $u_X(W) < u_X(T) < u_X(G)$ for any $X \in \{G, T, W\}$. However, those rankings might be different in a weakly coherent case, and we might for example have $u_W(W) < u_W(G) < u_W(T)$, $u_T(W) < u_T(T) < u_T(G)$ and $u_G(T) < u_G(W) < u_G(G)$. The preference relation $\omega_{\mathcal{D}}$ resulting from those rankings is indeed transitive and gives the ranking $G \succ T \succ W$ just as in the coherent case.

Note that beside the distortion introduced by experience there might be a good, interpretable reason for these differences between the rankings. In a direct comparison between Godzilla and Titanic for example, the consumer may prefer Godzilla as they consider it a less boring alternative, but when they make the comparison using Gone with the Wind as the reference point, choosing Godzilla might feel a little barbaric and even though that film is still preferable to a really boring Gone with the Wind it might suddenly seem a less appealing alternative than Titanic. In some situations, such as when the consumer is influenced by advertising that suggests Gone with the Wind should be taken as the reference point, it certainly feels plausible that even though $\omega_{\mathcal{D}}$ is transitive, the consumer might actually use u_W as a choice function and pick Titanic instead.

Note that we still do not know whether any weakly coherent measure $\mu_{\mathcal{D}}$ actually exists. However, before we can answer this question in the affirmative in theorem 2, we show in proposition 5 that if $\omega_{\mathcal{D}}$ is transitive, then we can easily give a utility function that represent it.

Proposition 5. *Let $\omega_{\mathcal{D}}$ and the weakly coherent $\mu_{\mathcal{D}}$ representing it be given. Then*

$$u_{\mathcal{D}}(x) = \lambda(\{y \in \mathcal{B} : x \succeq y\})$$

is a utility function representing $\omega_{\mathcal{D}}$.

Proof. As $\omega_{\mathcal{D}}$ is transitive $z_1 \succeq z_2$ implies that the lower contour sets $LCS_x = \{y \in \mathcal{B} : x \succeq y\}$ satisfy $LCS_{z_2} \subset LCS_{z_1}$. Moreover, if $z_1 \succ z_2$ then there exists some open ball $B \subset \mathcal{B}$ such that $B \subset LCS_{z_1}$ and $B \cap LCS_{z_2} = \emptyset$. Therefore, from the definition of measure λ on $\mathcal{B}$, $\lambda(LCS_{z_1}) > \lambda(LCS_{z_2})$ and as a consequence $u_{\mathcal{D}}$ represents $\omega_{\mathcal{D}}$.

□

Note that the characterisation given by proposition 5 is clearly not unique, as any function defined on the lower counter sets of the relation $\omega_{\mathcal{D}}$ that is strictly increasing with respect to the inclusion of those sets represents $\omega_{\mathcal{D}}$. However, the utility function given in this proposition is quite an interesting choice function in itself, since if $\mu_{\mathcal{D}}$ is not even weakly coherent, and $\omega_{\mathcal{D}}$ is not a preference relation, $u_{\mathcal{D}}$ still represents some preference relation that is consistent with $\omega_{\mathcal{D}}$ on any subset of $\mathcal{B}$ such that $\omega_{\mathcal{D}}$ restricted to this subset is transitive. We denote the preferences represented by the utility function $u_{\mathcal{D}}$ as $\omega_{\mathcal{B}}$. This preference relation is also quite intuitive, as this $u_{\mathcal{D}}$ essentially states that x is better than y if and only if x is preferred (with respect to $\omega_{\mathcal{D}}$) to a larger subset of alternative choices than y, no matter the direct comparison between x and y; this can be understood as a majority vote by the indirect utility functions u_z. As such, this is another case of indirect comparison between alternatives, but this time it is one that always results in a well–defined, single preference relation on the whole $\mathcal{B}$, so even though $\omega_{\mathcal{D}}$ is conceptually easier and intuitively more viable, $\omega_{\mathcal{B}}$ is actually more general and has nicer properties when the measure is not weakly coherent.

Note, however, that $\omega_{\mathcal{B}}$ can be sensitive to changes in the set of alternative choices. Even though we assume that $\mathcal{B}$ is known and preferences are defined on the whole $\mathcal{B}$, there are actually many situations where this might not be the case. In these cases, we might actually want to use $\omega_{\mathcal{B}}$ instead of $\omega_{\mathcal{D}}$, in order to account for the menu effects. To visualise this better, consider the further generalisation of $\omega_{\mathcal{B}}$ provided next by definition 13 and example 4.

Definition 13. *Let $A \subset \mathcal{B}$ be such that $\lambda(A) > 0$. We denote by $\succeq_A$ (also, ω_A) the menu dependent preference relation, defined as $x \succeq_A y \iff \lambda(\{z \in A : x \succ z\}) \geq \lambda(\{z \in A : y \succ z\})$.*

Example 4. Consider the example of a consumer who is faced with a choice of a new washing machine and is very much unaware of any of the sophisticated features that might be in new models. If this consumer goes first to some old–fashioned shop with a very limited assortment of older and less sophisticated models, we might reasonably expect that it would not cross the consumer's mind that there might be some very different alternative choices. It would therefore be reasonable to assume that this consumer might evaluate all the alternatives from set A, which consists of all the washing machines that they can imagine at the moment, which might be the set of washing machines available in the shop. Accordingly, the consumer uses ω_A as the preference relation when choosing, and so we assume that they choose washing machine 1. However, the consumer later enters a larger, more upmarket shop that has a lot of new, expensive, sophisticated machines in addition to everything that was in the previous shop an this enriches their knowledge of the possibilities available to an extended set A' such that $A \subset A'$. Even if those new, more sophisticated machines are irrelevant alternatives (but $\mu_{\mathcal{D}}$ is not weakly coherent) it might be that the consumer now chooses machine 2 that was also present in the previous shop but was not chosen, just because machine 1 might very well be perceived as worse in direct comparison with the new and more sophisticated machines, while machine 2 might seem better than them. This situation and the choice of the consumer as described in this example seem plausible, and can be accounted for by using ω_A.

We are now ready to go back to the question of whether the weakly coherent measures $\mu_{\mathcal{D}}$ actually exist, which is the main result of this section, given by theorem 2 below. It is preceded by supplementary definitions 14-16 and lemma 4, which are mostly of technical importance

Definition 14. *We denote by* $\mathrm{Diag}(\omega), \mathrm{Diag}^+(\omega), \mathrm{Diag}_-(\omega) \subset \mathcal{B} \times \mathcal{B}$ *sets of respectively diagonal, upper diagonal and lower diagonal elements of relation* ω*, that is*

$$\mathrm{Diag}(\omega) = \{(x, y) \in \mathcal{B} \times \mathcal{B} : x \sim_\omega y\},$$

$$\mathrm{Diag}^+(\omega) = \{(x, y) \in \mathcal{B} \times \mathcal{B} : x \succ_\omega y\},$$

$$\mathrm{Diag}_-(\omega) = \{(x, y) \in \mathcal{B} \times \mathcal{B} : x \prec_\omega y\}.$$

Definition 15. *Let* $\mu_{\mathcal{D}}$ *be given. We say that a measure* $\mu'_{\mathcal{D}}$ *is obtained from* $\mu_{\mathcal{D}}$ *by a disturbance* (μ', w) *if* μ' *is a probability measure defined on* $\Omega(\mathcal{D})$*, function* $w' : \mathcal{B}^2 \to [0, 1]$ *satisfy* $w'(x, y) = w'(y, x)$ *and*

$$\mu'_{\mathcal{D}}([x \succ y]) = (1 - w'(x, y))\mu_{\mathcal{D}}([x \succ y]) + w'(x, y)\mu'([x \succ y]).$$

Definition 16. *Let* $\mu_{\mathcal{D}}$ *be given. We say that disturbance* (μ', w) *does not disturb the diagonal, if and only if for* $A = \mathrm{supp}(w')$ *we have*

1. $A \cap \mathrm{Diag}(\omega_{\mathcal{D}}) = \emptyset$,

2. $(x, y) \in A \cap \mathrm{Diag}^+(\omega_{\mathcal{D}}) \implies \mu'([x \succ y]) \geq \frac{1}{2}$, *with equality only for* $w'(x, y) < 1$,

3. $(x, y) \in A \cap \mathrm{Diag}_-(\omega_{\mathcal{D}}) \implies \mu'([x \succ y]) \leq \frac{1}{2}$, *with equality only for* $w'(x, y) < 1$.

If this is not the case, we say that (μ', w') *disturb the diagonal.*

Lemma 4. *Let* $\mu_{\mathcal{D}}$ *be given and* $\mu'_{\mathcal{D}}$ *be obtained from* $\mu_{\mathcal{D}}$ *by a disturbance* (μ', w') *that does not disturb the diagonal. Then* $\mu'_{\mathcal{D}}$ *also represents* $\omega_{\mathcal{D}}$*.*

Proof. First, let $\mu'_{\mathcal{D}}$ be obtained from $\mu_{\mathcal{D}}$ without disturbing the diagonal and denote by $\omega'_{\mathcal{D}}$ (or $\succeq_{\mathcal{D}'}$) the relation given by definition 9 applied to $\mu'_{\mathcal{D}}$. Fix an arbitrary $x \in \mathcal{B}$. Following definition 16 we have $A \cap \{y \in \mathcal{B} : y \sim_{\mathcal{D}} x\} = \emptyset$. Therefore $x \sim_{\mathcal{D}} y \implies x \sim_{\mathcal{D}'} y$. Now let $y \in \mathcal{B}$ be such that $y \succ_{\mathcal{D}} x$. If $(y, x) \notin \text{supp}(w')$ then obviously $y \succ_{\mathcal{D}'} x$, and so therefore assume that $(y, x) \in \text{supp}(w')$. Now by definition of a disturbance

$$\mu'_{\mathcal{D}}([y \succ x]) = (1 - w'(y, x))\mu_{\mathcal{D}}([y \succ x]) + w'(y, x)\mu'([y \succ x]).$$

By assumption $y \succ_{\mathcal{D}} x$ we have $\mu_{\mathcal{D}}([y \succ x]) > \frac{1}{2}$. Moreover following definition 16 we have $\mu'([y \succ x]) \geq \frac{1}{2}$. Therefore $\mu'_{\mathcal{D}}([y \succ x]) > \frac{1}{2}$ and $y \succ_{\mathcal{D}'} x$. As the case with $x \succ_{\mathcal{D}} y$ is symmetric to this one, $\omega_{\mathcal{D}} = \omega'_{\mathcal{D}}$ and therefore $\mu'_{\mathcal{D}}$ also represents $\omega_{\mathcal{D}}$.

□

Theorem 2. *Let $\mathcal{D}$ be fixed. For any given $\omega \in \Omega(\mathcal{D})$ there exists some weakly coherent measure $\mu_{\mathcal{D}}$ representing ω. Moreover, for some fixed $\omega_{\mathcal{D}} \in \Omega(\mathcal{D})$, there is a $\mu_{\mathcal{D}}$ representing $\omega_{\mathcal{D}}$ that can be represented as*

$$\mu_{\mathcal{D}}([x \succeq y]) = \sum_{i=1}^{n} w_i(x, y)\mu_i([x \succ y]) + (1 - \sum_{i=1}^{n} w_i(x, y))\mu_*([x \succeq y]),$$

where (μ_i, w_i) are disturbances that do not disrupt the diagonal that for all $x, y \in \mathcal{B}$ satisfy $\sum_{i=1}^{n} w_i(x, y) \leq 1$ and μ_* is a coherent measure on Ω representing $\omega_{\mathcal{D}}$.

Proof. Due to corollary 1 we can define $\mu_{\mathcal{D}}$ only on the sets of the form $[x \succ y]$, knowing that we can always define other conditional measures in such a way, that there exists some μ defined on the whole Borel σ-field such that $\mu_{\mathcal{D}}$ is a conditional measure for μ when conditioned of $\mathcal{D}$. Therefore, we can restrict our attention to values of $\mu_{\mathcal{D}}$ on the sets $[x \succ y]$ only.

As $\omega_{\mathcal{D}} \in \Omega(\mathcal{D})$ there is a continuous utility function that represents it. Let u be this utility function, and denote by x^*, y_* some maximum and minimum elements for relation $\omega_{\mathcal{D}}$. As $\mathcal{B}$ is compact and $\omega_{\mathcal{D}}$ is continuous, such x^*, y_* exist. Define $\mu_*([x \succ y]) = \frac{1}{2} + \frac{u(x)-u(y)}{2(u(x^*)-u(y_*))}$. Clearly this $\mu_*([x \succeq y]) \geq \frac{1}{2} \iff u(x) \geq u(y)$, and therefore it represents $\omega_{\mathcal{D}}$ on Ω. Moreover, for any $z \in \mathcal{B}$ we have $\mu_*([x \succeq z]) \geq \mu_*([y \succeq z]) \iff u(x) \geq u(y)$ and therefore this μ_* is coherent. However, it cannot represent $\omega_{\mathcal{D}}$ on $\Omega(\mathcal{D})$ as it is not restricted to $\Omega(\mathcal{D})$, so for $d_1 \succ d_2 \in \mathcal{K}_{\mathcal{D}}$ does not imply $\mu_*([d_1 \succ d_2]) = 1$ unless $d_1 \sim_{\mathcal{D}} x^*$ and $d_2 \sim_{\mathcal{D}} y_*$. Note however, that from definition 9 we have $d_1 \succ d_2 \in \mathcal{K}_{\mathcal{D}} \iff d_1 \succ_{\mathcal{D}} d_2$, and therefore $\mu_*([d_1 \succ d_2]) > \frac{1}{2}$.

By lemma 4 if we disturb μ_* without disturbing the diagonal, the disturbed measure also represents $\omega_{\mathcal{D}}$. We now show that there is a sequence $(\mu_i, w_i)_{i=1}^n$ of disturbances that does not disturb the diagonal, such that $(1-\sum_{i=1}^n w_i(x,y))\mu_*([x \succ y]) + \sum_{i=1}^n w(x,y)\mu_i([x \succ y])$ is equal to 0 whenever $y \succeq x \in \mathcal{K}_{\mathcal{D}}$. Due to axiom 6 we can assume without loss of generality that $\mathcal{K}_{\mathcal{D}}$ consists of strict preference relations only and we denote all known relations as $\mathcal{K}_{\mathcal{D}} = \{x_i \succ y_i : i \leq n\}$.

For all i fix some pairwise disjoint $B_i = B((x_i, y_i), r_i)\mathrm{Diag}^+(\omega_{\mathcal{D}})$ and define

$$w_i(x,y) = \max\left\{1 - \frac{d((x,y),(x_i,y_i))}{r_i}, 1 - \frac{d((x,y),(y_i,x_i))}{r_i}, 0\right\},$$

$$u_i(x) = \begin{cases} 1 & \text{if } u(x) > u(x_i), \\ 0 & \text{if } u(x) < u(y_i), \\ \frac{u(x)-u(y_i)}{u(x_i)-u(y_i)} & \text{otherwise.} \end{cases}$$

It suffices to take $\mu_i([x \succ y]) = \frac{1}{2} + \frac{u(x)+u(y)}{2}$. By construction each disturbance (μ_i, w_i) does not disturb the diagonal and as a result $\mu_{\mathcal{D}}([x \succ y]) = (1 - \sum_{i=1}^n w_i(x,y))\mu_*([x \succ y]) + \sum_{i=1}^n w(x,y)\mu_i([x \succ y])$ represents $\omega_{\mathcal{D}}$. Moreover as $w_i(x_i, y_i) = 1$ and $\mu_i([x_i \succ y_i]) = 1$ so μ is restricted to $\Omega(\mathcal{D})$. As additionally $\mu_{\mathcal{D}}$ is of the requested form, the proof is finished.

□

Theorem 2 shows something more than just the existence of a weakly coherent $\mu_{\mathcal{D}}$. First, it shows that by choosing $\mu_{\mathcal{D}}$ we can represent any $\omega \in \Omega(\mathcal{D})$ as conditional preferences, meaning any preferences in $\Omega(\mathcal{D})$ are permissible. Secondly, and more importantly, it also shows what weakly coherent measures actually look like; they are created as a coherent measure that is locally disturbed by some other measures in a way that does not disturb the diagonal. We already noted in the comments after proposition 3 that its proof hints that the reason for the non–existence of coherent measures is the disturbance introduced by experience. Indeed, in the proof of theorem 2 we explicitly constructed $\mu_{\mathcal{D}}$ as a coherent measure with disturbances introduced by each known relation between the alternative choices. This observation is formalised in the corollary 2 as follows.

Corollary 2. *Let μ_* be a coherent measure on Ω representing $\omega_{\mathcal{D}}$ and $\mu_{\mathcal{D}}$ be given as $\mu_{\mathcal{D}}([x \succ y]) = (1 - w'(x,y))\mu_*([x \succ y]) + w'(x,y)\mu'(x,y)$ for some disturbance μ' that does not disturb the diagonal. Then for all pairs $x, y \in \mathcal{D}$ we* *have* $w'(x_i, y_i) = 1$*. Moreover, let* U *be an arbitrary open subset of* $\mathcal{B}^2$ *such that* $x, y \in \mathcal{D} \implies (x, y) \in U$*. Then for any coherent* μ_* *that represents* $\omega_{\mathcal{D}}$ *on* Ω *there exists a disturbance* (μ', w') *such that* $\mathrm{supp}(w') \subset U$ *and* $(1 - w'(x, y))\mu_*([x \succ y]) + w'(x, y)\mu'(x, y)$ *represents* $\omega_{\mathcal{D}}$ *on* $\Omega(\mathcal{D})$*.*

Proof. This follows straight from the construction of

$$\mu_{\mathcal{D}}([x \succ y]) = (1 - \sum_{i=1}^{n} w_i(x, y))\mu_*([x \succ y]) + \sum_{i=1}^{n} w_i(x, y)\mu_i([x \succ y])$$

in the proof of theorem 2.

□

This corollary tells us firstly that any coherent measure μ_* must be disturbed in all points in $\mathcal{B}^2$ on which the relation is known and secondly that those points are the only ones in which the disturbance is really necessary in order to obtain $\mu_{\mathcal{D}}$ that represents $\omega_{\mathcal{D}}$ on $\Omega(\mathcal{D})$ from some coherent μ_* on Ω. Of course, continuity means, that the disturbance must spill over to the neighbourhood of those points. Note, however, that nothing stops us adding some completely unnecessary disturbance to $\mu_{\mathcal{D}}$ if we feel like it, as long as it does not disturb the diagonal.

5 Experimental preferences

As noted already, in the setting with unknown preferences, there are two easily identifiable motivations for the consumer. The first is the utility obtained from immediate consumption, and the other is the chance to obtain better knowledge about their own preferences in order to make better, more informed choices in the future; the sole purpose of exploration of own preferences can also be enjoyable for some. As experimentation is primarily a dynamic feature of preference discovery, we do not study it in the same level of detail as conditional preferences. However, we define the preference relation that is responsible for the experimental motivation of the consumer, and we explore its properties a little.

The first intuition for how to define experimental preferences is to infer them from some expected discounted utility formula by ignoring the current period. However, in the ordinal case that we are in, it is simply not possible to do this in any natural way. The second intuition, and the one we follow, is to define somehow how much information is gained by adding a given alternative x to $\mathcal{D}$.

Definition 17. *Let $\mu_{\mathcal{D}}$ be given. We call relation $\succeq_E$ (also denoted ω_E) defined by $x \succeq_E y \iff E_\mu[\mu(\Omega(\mathcal{D} \cup \{x\})] \leq E_\mu[\mu(\Omega(\mathcal{D} \cup \{y\})]$ the experimental preference relation of the consumer (conditional on $\mathcal{D}$). Accordingly, for a given preference relation $\succeq_E$ we say that μ such that $E_\mu[\mu(\Omega(\mathcal{D} \cup \{x\})] \leq E_\mu[\mu(\Omega(\mathcal{D} \cup \{y\})]$ represents $\succeq_E$ as the experimental preference relation of the consumer.*

For brevity, we call ω_E experimental preferences, whereas we simply say that μ represents ω_E. To give a better intuition about definition 17 we give the following explanation. Given all the knowledge the consumer already has from knowing $\mathcal{D}$, alternative x is more informative than y if adding x to $\mathcal{D}$ results in $\Omega(\mathcal{D}) \cup \{x\}$ being a smaller subset of Ω than the $\Omega(\mathcal{D}) \cup \{y\}$ that we would obtain by adding y. A natural way to measure the size of the subsets of Ω is by using the measure μ, therefore defining x as more informative than y if $\mu(\Omega(\mathcal{D}) \cup \{x\}) < \mu(\Omega(\mathcal{D}) \cup \{y\})$. However ex ante, the consumer does not know for certain, how large $\Omega(\mathcal{D}) \cup \{x\}$ will turn out to be. They only know the size of $\Omega(\mathcal{D} \cup \{x\})$ conditionally on the relations between x and elements of $\mathcal{D}$ and the probabilities of those exact relations, which are also given by μ. Therefore, ex ante the best they can do is to take the expected value of $\mu(\Omega(\mathcal{D}) \cup \{x\})$ with respect to measure μ, thereby obtaining $E_\mu[\mu(\Omega(\mathcal{D} \cup \{x\})]$. Proposition 6 states some basic facts about ω_E.

Proposition 6. *Experimental preferences ω_E are always complete, transitive, continuous and reflexive. Moreover let $\mathcal{K}_{\mathcal{D}} = \{x_1 \succeq \cdots \succeq x_n\}$. Then the utility function $u_E(x) = 1 - \sum_{i=1}^{n-1} \mu_{\mathcal{D}}^2([x_i \succ x \succ x_{i+1}]) - \mu_{\mathcal{D}}^2([x \succ x_1]) - \mu_{\mathcal{D}}^2([x_n \succ x])$ represents ω_E.*

Proof. As u_E as defined in the statement of the theorem is continuous, the second part of the theorem implies the first part. Therefore we only need to prove that u_E represents ω_E. From definition 2 we have the following

$$E_\mu[\mu(\Omega(\mathcal{D} \cup \{x\})] = \mu([\mathcal{K}_{\mathcal{D}} \cup \{x_n \succ x\}])\mu_{\mathcal{D}}([x_n \succ x]) + \mu([\mathcal{K}_{\mathcal{D}} \cup \{x \succ x_1\}])\mu_{\mathcal{D}}([x \succ x_1]) +$$

$$+ \sum_{i=1}^{n-1} \mu([\mathcal{K}_{\mathcal{D}} \cup \{x_i \succ x, x \succ x_i + 1\}])\mu_{\mathcal{D}}([x_i \succ x \succ x_{i+1}]) =$$

$$= \mu_{\mathcal{D}}^2([x_n \succ x])\mu(\Omega(\mathcal{D})) + \mu_{\mathcal{D}}^2([x \succ x_1])\mu(\Omega(\mathcal{D})) + \sum_{i=1}^{n-1} \mu_{\mathcal{D}}^2([x_i \succ x \succ x_{i+1}])\mu(\Omega(\mathcal{D})),$$

and therefore

$$x \succeq_E y \iff (1 - u_E(x))\mu(\Omega(\mathcal{D})) \leq (1 - u_E(y))\mu(\Omega(\mathcal{D})) \iff u_E(x) \geq u_E(y).$$

□

As shown by proposition 6 it is trivial that ω_E is transitive, which was so problematic in the case of $\omega_{\mathcal{D}}$, and just as trivial is obtaining the natural utility representation of ω_E. To understand better this utility representation, consider the characterisation of the optimal element with respect to ω_E. Note that this optimal element exists, since u_E is continuous and $\mathcal{B}$ is compact. Denoting $p_i(x) = \mu_{\mathcal{D}}([x_i \succ x \succ x_{i+1}])$, $p_0(x) = \mu_{\mathcal{D}}([x \succ x_1])$ and $p_n(x) = \mu_{\mathcal{D}}([x_n \succ x])$, we have $\max_{x\in\mathcal{B}} \sum_{i=0}^n p_i^2$ under the condition that $\sum_{i=0}^n p_i = 1$. Clearly, if this element exists in $\mathcal{B}$, the solution to this optimisation problem is $x^* \in \mathcal{B}$ such that $p_i(x) = \frac{1}{n+1}$. This condition is easily interpretable, as this x^* is precisely the element, that the consumer has no knowledge about at all, or has the least knowledge about if there is no element with all $p_i = \frac{1}{n+1}$ in terms of its preference ranking against the known alternatives; this lack of knowledge is reflected by each position in the resulting ranking being just as probable for x^*.

To simplify notation, we keep the notation that we introduced in the proof of proposition 6 throughout this whole section, so $p_i = \mu_{\mathcal{D}}([x_i \succ x \succ x_{i+1}])$, $p_0(x) = \mu_{\mathcal{D}}([x \succ x_1])$ and $p_n(x) = \mu_{\mathcal{D}}([x_n \succ x])$. One obvious consequence of proposition 6 is that to determine experimental preferences completely, all that we need are the indirect utility functions u_i for $x_i \in \mathcal{D}$, instead of the whole measure. This obvious observation is formally stated in corollary 3.

Corollary 3. *Let* $\mathcal{D} = \{x_1, \ldots, x_n\}$, $\forall_{i_1<i_2} x_{i_1} \succ x_{i_2} \in \mathcal{K}$ *and assume we are given* $\mu_{\mathcal{D}}, \mu'_{\mathcal{D}}$ *such that indirect utility functions* $u_i(x) = \mu_{\mathcal{D}}([x \succ x_i])$ *and* $u'_i(x) = \mu'_{\mathcal{D}}([x \succ x_i])$ *are equal. Then experimental preferences* ω_E, ω'_E *(experimental preferences with respect to* $\mu_{\mathcal{D}}, \mu'_{\mathcal{D}}$ *respectively) are equal.*

Proof. To see this, it suffices to show that u_E as defined in the proposition 6 can be written as a function of $u_i(x)$ only. Using the notation introduced in the proof of proposition 6, we know that $u_E(x) = 1 - \sum_{i=0}^n p_i^2(x)$, whereas $p_0 = u_1$, $p_n = 1 - u_n$ and for $1 \leq i \leq n-1$ we have $p_i = u_{i+1} - u_i$.

□

Note that conversely to the assignment in the proof of corollary 3 we can identify u_i given p_i, that is $u_i = \sum_{j=0}^{i-1} p_j$.

The next natural question about ω_E is what conditions some element $\omega \in \Omega$ must satisfy so that there is some $\mu_{\mathcal{D}}$ for which $\omega = \omega_E$. Some properties are obvious, as shown by proposition 7.

Proposition 7. *Let ω be such that there exists $\mu_{\mathcal{D}}$ for which $\omega = \omega_E$. Then for all $x, y \in \mathcal{D}$ we have $x \sim y$ and for any $x \in \mathcal{D}, y \in \mathcal{B} \setminus \mathcal{D}$ we have $y \succ x$.*

Proof. It follows trivially from the fact that $\mu(\Omega(\mathcal{D} \cup \{x\})) = \mu(\Omega(\mathcal{D}))$ for $x \in \mathcal{D}$ and that $\mu(\Omega(\mathcal{D} \cup \{y\})) < \mu(\Omega(\mathcal{D}))$ for any $y \notin \mathcal{D}$, where the second fact follows from axiom 4.

□

Proposition 7 shows the necessary conditions on ω for it to be able to be represented as experimental preferences. The question of whether this condition is also sufficient is hard to answer in the general case, and is beyond the scope of this article.

6 Conclusions

In this article, we provided the mathematical foundations for considering a consumer with imperfect knowledge of their own preferences. We proposed that this consumer should be considered not as equipped with some preference relation, but rather with a probabilistic measure on a space of all the possible rational preference relations. To consider these measures, we have defined a topological structure on this space that we perceive to be extremely natural and connected to the information structure that is available to the consumer. In theorem 1 we provide a very important tool that allows us to easily define these measures. Especially significant aspect of the tool provided by theorem 1 is that it emphasises the importance of indirect learning as a determining factor in measure definition.

Based on these measures, we define and study three choice procedures that correspond to the perception of their own real preferences by the consumer under different assumptions, namely indirect preferences, conditional preferences and menu-dependent preferences. Especially important are conditional preferences, and we feel this relation is the most appropriate under normal circumstances, meaning when menu is given and measure is weakly coherent. Finally, we define and discuss experimental preferences of the consumer, that describe the ranking of alternatives based on the perceived value of information that the consumption of any given alternative would bring.

We feel that this work opens up many possibilities for further research in preference discovery. One obvious route for extension is to consider a dynamic case, looking at paths of consumption generated by some choice functions, and how the conditional and experimental preferences evolve along those paths. Such an extension seems very natural, especially as the question of how well the consumers discover their preferences is at the very core of this theory.

Another area for further research is to extend the framework from this article to a richer learning setting. In the current work, the learning of the consumer is very simplistic as there are no errors in perception, the perception is perfect and the consumer is Bayesian. Such an extension would allow several interesting questions to be answered, since for example the assumption of non-Bayesian updating can result in a lack of convergence as shown by Epstein et al. (2010), and as a result there might be imperfect preference discovery even in situations in which it would usually occur.

Finally, the current work leaves unanswered the question of whether any empirical verification of preference discovery is possible. Note that the presence of an exogenous similarity metric in our framework means there are two subjective elements; the first is the measure that describes the perception of the consumer, and the second is the similarity metric. In general, it is certainly possible that any behaviour of the consumer could be accommodated by simply changing the similarity metric. This need for dual identification means the question of the verifiable conditions of this model seems very interesting.

Acknowledgements

The study was funded by the National Science Centre in Poland (Grant No. 2020/37/N/HS4/03367).

7 References

Ariely, D. and Wallsten, T.S. (1995). Seeking subjective dominance in multidimensional space: an exploration of the asymmetric dominance effect. Organizational

Behavior and Human Decision Processes 63(3), p. 223–232.

Ariely, D., Loewenstein, G. and Prelec, D. (2006). Tom Sawyer and the construction of value. Journal of Economic Behavior & Organization 60(1), p. 1-10.

Ash, R.B. (1999). Probability and Measure Theory (2nd ed.). Academic Press, p. 19.

Bettman, J.R., Luce, M.F. and Payne, J.W. (1998). Constructive consumer choice processes. Journal of Consumer Research 25(3), p. 187–217.

Bleichrodt, H. and Pinto Prades, J.L. (2009). New evidence of preference reversals in health utility measurement. Health Economics 18(6), p. 713–726.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2012a). Salience in experimental tests of the endowment effect. American Economic Review 102(3), p. 47-52.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2012b). Salience theory of choice under risk. The Quarterly Journal of Economics 127(3), p. 1243–1285.

Braga, J. and Starmer, C. (2005). Preference anomalies, preference elicitation and the discovered preference hypothesis. Environmental & Resource Economics 32, p. 55-89.

Braga, J., Humphrey, S.J. and Starmer, C. (2009). Market experience eliminates some anomalies—and creates new ones. European Economic Review 53(4), p. 401-416.

Bruni, L. and Sugden, R. (2007). The road not taken: how psychology was removed from economics, and how it might be brought back. The Economic Journal 117, p. 146-173.

Butler, D.J. and Loomes, G.C. (2007). Imprecision as an account of the preference reversal phenomenon. American Economic Review 97(1), p. 277-297.

Carlsson, F., Mørkbak, M.R. and Olsen, S.B. (2012). The first time is the hardest: a test of ordering effects in choice experiments. Journal of Choice Modelling 5(2), p. 19-37.

Cooke, K. (2017). Preference discovery and experimentation. Theoretical Economics 12, p. 1307-1348.

Coupey, E. (1994). Restructuring: Constructive processing of information displays in consumer choice. Journal of Consumer Research 21(1), p. 83-99.

Cox, J.C. and Grether, D.M. (1996). The preference reversal phenomenon: response mode, markets and incentives. Econ Theory 7, p. 381–405.

Czajkowski, M., Hanley, N. and LaRiviere, J. (2015). The effects of experience on preferences: theory and empirics for environmental public goods. American Journal of Agricultural Economics 97, p. 333-351.

Day, B., Bateman, I.J., Carson, R.T., Dupont, D., Louviere, J.J, Morimoto, S., Scarpa, R. and Wang, P. (2012). Ordering effects and choice set awareness in repeat-response stated preference studies. Journal of Environmental Economics and Management 63(1), p. 73-91.

Debreu, G. (1954). Representation of a preference ordering by a numerical function. In: Thrall, M., Davis, R.C. and Coombs, C.H. (eds.), Decision Processes. John Wiley and Sons, New York, 159-165.

Debreu, G. (1959). Theory of value. An axiomatic analysis of economic equilibrium. New Haven and London, Yale University Press.

Dekel, E., Lipman, B.L. and Rustichini, A. (2001). Representing preferences with a unique subjective state space. Econometrica 69(4).

Delaney, J.J., Jacobson, S.A. and Moenig, T.P. (2014). Discovered preferences for risky and non-risky goods. Department of Economics Working Papers 2014-02, Department of Economics, Williams College.

Delaney, J.J., Jacobson, S.A. and Moenig, T.P. (2019). Preference discovery. Experimental Economics.

DellaVigna, S. (2009). Psychology and economics: evidence from the field. Journal of Economic Literature 47(2), p. 315-72.

Enderton, H.B. (2001). A Mathematical Introduction to logic (2nd ed.). Academic Press.

Engelmann, D. and Hollard, G. (2010). Reconsidering the effect of market experience on the "endowment effect". Econometrica 78, p. 2005-2019.

Epstein, L.G., Noor J. and Sandroni A. (2010). Non-bayesian learning. The B.E. Journal of Theoretical Economics 10(1).

Gilboa, I. and Schmeidler, D. (2001). A theory of case-based decisions. Cambridge: Cambridge University Press.

Glöckner, A., Betsch, T. and Schindler, N. (2010). Coherence shifts in probabilistic inference tasks. Journal of Behavioral Decision Making 23(5), p. 439–462.

Grether, D. and Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. The American Economic Review 69(4), p. 623-638.

Guo, L. (2021). Testing the role of contextual deliberation in the compromise effect. Management Science 0(0), p. 1-19.

Hoeffler, S. and Ariely, D. (1999). Constructing stable preferences: a look into dimensions of experience and their impact on preference stability. Journal of Consumer Psychology 8(2), p. 113-139.

Huang, S., Su, X., Hu, Y., Mahadevan, S. and Deng Y. (2014). A new decision-making method by incomplete preferences based on evidence distance. Knowledge-Based Systems 56, p. 264-272.

Humphrey, S.J., Lindsay, L. and Starmer, C. (2017). Consumption experience, choice experience and the endowment effect. J Econ Sci Assoc. 3(2), p. 109-120.

Kahneman, D., Knetsch, J.L. and Thaler, R.H. (1991). Anomalies: the endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives 5(1), p. 193-206.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica 47(2), p. 263-91.

Kannai, Y. (1970). Continuity properties of the core of a market. Econometrica 38(6), p. 791–815.

Kingsley, D.C. and Brown, T.C. (2010). Preference uncertainty, preference learning, and paired comparison experiments. Land Economics 86(3), p. 530-544.

Kreps, D.M. (1979). A representation theorem for “preference for flexibility”. Econometrica 47(3), p. 565-577.

Kuilen, G. van de (2009). Subjective probability weighting and the discovered preference hypothesis. Theory and Decision 67, p. 1-22.

Lerner, J.S., Li, Y., Valdesolo, P. and Kassam K.S. (2015). Emotion and decision making. Annual Review of Psychology 66(1), p. 799-823.

Lichtenstein, S. and Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. Journal of Experimental Psychology 89, p. 46–55.

Lichtenstein, S. and Slovic, P. (1973). Response-induced reversals of preference in gambling: an extended replication in Las Vegas. Journal of Experimental Psychology 101(1), p. 16–20.

Lichtenstein, S. (eds.) and Slovic, P. (eds.) (2006). The construction of preference. Cambridge University Press.

Loewenstein, G. and Issacharoff, S. (1994). Source dependence in the valuation of objects. Journal of Behavioral Decision Making 7(3), p. 157–168.

Loomis, J. (2011). What’s to know about hypothetical bias in stated preference valuation studies?. Journal of Economic Surveys 25, p. 363-370.

Loomes, G. and Sugden, R. (1982). Regret theory: an alternative theory of rational choice under uncertainty. The Economic Journal 92(368), p. 805-824.
Loomes, G., Orr, S. and Sugden, R. (2009). Taste uncertainty and status quo effects in consumer choice. Journal of Risk and Uncertainty 39(2), p. 113–135.
Mandel, N. and Johnson, E.J. (2002). When web pages influence choice: effects of visual primes on experts and novices. Journal of Consumer Research 29(2), p. 235–245.
Nowlis, S.M. and Simonson, I. (1997). Attribute–task compatibility as a determinant of consumer preference reversals. Journal of Marketing Research 34(2), p. 205-218.
Payne, J.W., Bettman, J.R. and Schkade D.A. (1999). Measuring constructed preferences: towards a building code. Journal of Risk and Uncertainty 19, p. 243-270.
Piermont, E., Takeoka, N. and Teper, R. (2016). Learning the Krepsian state: exploration through consumption. Games and Economic Behavior 100, p. 69-94.
Plott C.R. (1996) Rational individual behaviour in markets and social choice processes: the discovered preference hypothesis. In: Arrow K.J., et al. (eds.). The rational foundations of economic behaviour: Proceedings of the IEA Conference held in Turin, Italy, IEA Conference Volume, no. 114. New York: St. Martin's Press; London: Macmillan Press in association with the International Economic Association, p. 225-250.
Plott, C.R. and Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. American Economic Review 95(3), p. 530-545.
Rizzo, M.J. and Whitman, G. (2018). Rationality as a process. Review of Behavioral Economics 5, p. 201-219.
Savage, L.J. (1954). The foundations of statistics. John Wiley & Sons.
Schkade, D.A. and Johnson E.J. (1989). Cognitive processes in preference reversals. Organizational Behavior and Human Decision Processes 44(2), p. 203-231.
Schkade, D.A. and Payne, J.W. (1994). How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. Journal of Environmental Economics and Management 26(1), p. 88-109.
Simon H.A. (1976). From substantive to procedural rationality. In: Kastelein

T.J., Kuipers S.K., Nijenhuis W.A., Wagenaar G.R. (eds.). 25 Years of Economic Theory. Springer, Boston, MA.

Montgomery, H. (1998). Decision making and action. In: Kofta, M., Weary, G. and Sedek, G. (eds.) Personal Control in Action. The Springer Series in Social Clinical Psychology. Springer, Boston, MA.

Simon, D., Krawczyk, D.C. and Holyoak, K.J. (2004). Construction of preferences by constraint satisfaction. Psychological Science 15(5), p. 331–336.

Simon, D., Krawczyk, D.C., Bleicher, A. and Holyoak, K.J. (2008). The transience of constructed preferences. J. Behav. Decis. Making 21, p. 1-14.

Simonson, I. and Tversky, A. (1992). Choice in context: tradeoff contrast and extremeness aversion. Journal of Marketing Research 29 (3): p. 281–295.

Slovic, P. (1995). The construction of preference. American Psychologist 50(5), p.364-371.

Stigler, G. and Becker, G. (1977). De gustibus non est disputandum. American Economic Review 67(2), p. 76-90.

Sugden, R. (2003). Reference-dependent subjective expected utility. Journal of Economic Theory 111(2), p. 173-191.

Trautmann, S.T., Vieider, D.M. and Wakker, P.P. (2011). Preference reversals for ambiguity aversion. Management Science 57(7), p. 1320-1333.

Tversky, A., Sattath, S. and Slovic, P. (1988). Contingent weighting in judgment and choice. Psychological Review 95, 371- 384.

Warren, C., McGraw, A.P. and Van Boven, L. (2011). Values and preferences: defining preference construction. WIREs Cogn Sci 2, p. 193-205.

Wilson, M.S. (2018). Rationality with preference discovery costs. Theory Decis 85, p. 233–251.