INE PAN Working Paper Series

Paper number 52

Preference discovery, consumer beliefs and preference reversal paradox

Marek Kapera

Warsaw, January 2024

wp@inepan.waw.pl

Preference discovery, consumer beliefs and preference reversal paradox

Marek Kapera*

March 18, 2024

Abstract

In this paper, I provide a theoretical explanation of the empirically observed connection between preference discovery and preference reversal paradox. While the connection between preference discovery and multiple paradoxes of consumer choice, most notably preference reversal, is well established in the empirical literature, existing models of taste uncertainty do not predict or accommodate this kind of behaviour. In order to do so, I formulate a very general model of how a taste uncertain consumer perceives their preferences. I propose to consider the consumer with a partial knowledge of their own preferences as forming probabilistic beliefs regarding their own preferences, which I represent as a probabilistic measure space. I define two conditional preference relations, which correspond to direct and indirect comparisons of alternatives under incomplete information regarding the consumer's own preferences. Using those relations, I show two possible sources

*Institute of Economics, Polish Academy of Sciences, 00-330 Warsaw, Nowy Swiat 72, Poland, kaperamar@gmail.com, +48 609 686 255, ORCID: 0000-0001-7028-5193

I thank my PhD supervisor Łukasz Woźny for his directions and suggestions. I am grateful to Paweł Dziewulski, Rui Silva, Karol Flores-Szwagrzak, Piotr Denderski, Rabah Amir, Miklós Pintér and Michał Lewandowski for helpful discussions. Comments from the audiences at the 2023 Africa Meeting of the Econometric Society, XXX European Workshop on Economic Theory, 17th Warsaw International Economic Meeting, 9th VOCAL Optimization Conference, Institute of Economics of the Polish Academy of Sciences, Warsaw School of Economics, University of Rijeka, Krakow university of Economics and Adam Mickiewicz University are also highly appreciated. The study was funded by NCN Grant UMO-2020/37/N/HS4/03367.

of preference reversal under taste uncertainty, that is the intransitivity of direct comparisons, and different preference orderings from those two preference relations, and I provide testable conditions for both of those. Finally, I show examples of other possible applications of the presented model.

Keywords: decision theory; learning through consumption; preference discovery; preference formation; preference reversal; taste uncertainty

JEL classification: D11, D83, D91

1 Introduction

Preference discovery, formulated by (Plott 1996) is a hypothesis which states that people do not have a intrinsic knowledge of their own preferences, but rather discover them in the process of consumption. In contrast to psychological theories of preference construction (for a comprehensive overview see Lichtenstein and Slovic 2006), preference discovery assumes that the consumer has some well-defined, real preferences, but those are ex-ante unknown to the consumer. Only after the consumer has experienced some alternatives, the relative preference ranking of those alternatives is revealed to them.

It is an intuitively valid hypothesis, and in the words of (Plott 1996) it is a hypothesis that most of the economists actually believe in but rarely state. This hypothesis also has important implications for economic theory, mainly because of the empirically established connection between preference discovery and the preference reversal paradox (e.g. Lichtenstein and Slovic 1971). It has been first observed by (Cox and Grether 1996), who has shown that preference reversals are less prevalent in repeated experiments with incentives to learn and based on this result, (Plott 1996) suggested that observed paradoxes of choice might occur because individuals in the experiments are asked to make choices which they rarely if ever perform in everyday life. As such, they might not know their own preferences and make mistakes.

Importance of preference discovery for paradoxes of choice is well supported by empirical studies. Not only the observed choices indeed stabilize in repeated experiments, as shown for example by (Kingsley and Brown 2010) and (Czajkowski, Hanley, and LaRiviere 2015), but a large body of evidence suggests that preference discovery can have far reaching consequences for observed behaviour. These include the aforementioned preference reversal (Cox and Grether 1996, Plott 1996, Butler and G. C. Loomes 2007), the WTP/WTA disparity (Plott and Zeiler 2005, Engelmann and Hollard 2010, Humphrey, Lindsay, and Starmer 2017) and the order effects in stated preference studies (Day et al. 2012, Carlsson, Mørkbak, and Olsen 2012). The results obtained by (Kuilen 2009) even suggest that preference discovery can 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, to the best of my knowledge there is no theoretical explanation whatsoever as to why any of this should be the case. The two explanations of preference reversal that are dominant in the literature are firstly, that of (Tversky, Sattath, and Slovic 1988) which states that different procedures are applied in choice and valuation tasks; and secondly, the reference dependent model of (Sugden 2003). Neither of those explain the empirical connection between taste uncertainty and observed reversals. Taste uncertainty is present in the economic literature at least since the contribution of (Kreps 1979), with two recent extensions of (Kreps 1979) model by (Piermont, Takeoka, and Teper 2016) and (Cooke 2017), but none of those models either predict or accommodate preference reversals. As such, the vague idea introduced by (Plott 1996) that individuals make mistakes when faced with a new decision problem is still the only theoretical justification of the link between taste uncertainty and observed paradoxes of choice. However we do not have any idea what is the nature of those mistakes and why they occur. If no new taste information is acquired in between, it is unclear why taste uncertainty should make consumer choices to reverse.

In order to fill this gap, I do not consider consumer choice itself, but rather the beliefs of the consumer regarding their own preferences. In the contemporary literature on taste uncertainty, it is assumed that the consumer comes equipped with interim preferences, meaning a preference relations which represent the consumer choices conditionally on the information available to the consumer. However, this assumption explicitly excludes the possibility of preference reversals. Therefore, I replace this assumption and consider the consumer as forming probabilistic beliefs regarding their own taste.

I consider two possible explanations for the observed link between taste uncertainty and preference reversal. Firstly, it is possible that the expectations of the consumer regarding their own taste are inconsistent, meaning that the beliefs of the consumer are such that preference relation defined by what the consumer expects the real relation between alternatives to be is intransitive. Secondly, the consumer may apply a different procedure to evaluate the alternatives between choice and valuation tasks. More precisely, in choice task the consumer is asked to directly compare the alternatives to one another, whereas in the valuation task the comparison is indirect. It is possible that under perfect information preference rankings obtained using both of these procedures coincide, but it does not have to be the case under taste uncertainty. For both of those hypotheses, I provide the property of the system of consumer beliefs which is equivalent to the possibility of preference reversal occurring in this way.

The structure of the article is as follows. All of the basic elements and definitions of the model are given in section 2. In section 3 I define the system of beliefs of the consumer and provide the theorem which allows for its identification. In the same section I also define the expected preferences of the consumer, together with a result which shows that for any preference relation which agrees with what the consumer has already learned regarding their taste, there exists a system of beliefs such that this preference relation is the expected preference relation of this consumer.

I consider the connection between taste uncertainty and preference reversal in section 4. For both of the hypotheses regarding the connection between taste uncertainty and preference reversal, my results in this section provide easily testable properties of the system of beliefs of the consumer which are equivalent to the possibility of preference reversals to occur in this way.

Even though the connection between preference discovery and preference reversal is the main focus of this article, my model of consumer beliefs is very general and can be treated as a very general language in which to consider a taste uncertain consumer and section 5 provides some applications and extensions of the model.

I conclude in section 6 with a discussion of my results and their connection to the existing literature. I keep short proofs in the main text, but the longer ones are in the appendix.

2 Model elementaries

There are three basic elements of the model, namely the set of alternative choices $\mathcal{X}$, the set of possible preferences Ω and the set of alternatives known to the consumer $\mathcal{K}$.

Set of alternative choices represents the consumption alternatives that the consumer evaluates. Possible preferences are all the binary relations on the set of alternative choices that might be the real preferences of the consumer. One element of this set is the real preference relation of the consumer and I denote this element by ω^*.

It is ex-ante unknown to the consumer which element of Ω is their real preference relation, but it is partially revealed after consumption. Set of the alternatives know to the consumer represents the alternatives which the consumer has already consumed and as such, they know the real preference ranking of those alternatives, meaning that the consumer knows the relation ω^* restricted to the subset of alternatives in $\mathcal{K}$.

In this section I introduce those three elements in more detail and provide the necessary mathematical structure. Additionally, I define some notation and operations to work with incomplete preferences, which are an important tool in the model.

Objects of choice

Let $(\mathcal{X}, d)$ be a metric space, such that in the topology induced by metric d the set $\mathcal{X}$ is compact and connected. Generic elements of $\mathcal{X}$ are denoted by x, y, z. Elements of $\mathcal{X}$ are the objects of choice, and the interpretation of metric d is as a measure of similarity of alternatives, meaning that if $d(x,y) < d(x,z)$ then I interpret x as more similar[1] to y than to z. Abusing notation a little, I also denote by d a product metric on $\mathcal{X} \times \mathcal{X}$ given by $d((x_1,y_1),(x_2,y_2)) = \sqrt{d(x_1,x_2)^2 + d(y_1,y_2)^2}$. Open balls in $\mathcal{X}$ with the centre at x and the radius r are given by $B(x,r)$.

[1]Technically d measures dissimilarity of the alternatives, but I call it a measure of similarity nevertheless.

Possible preferences

The set of possible preferences, denoted by Ω, is defined as a set of all binary relations on $\mathcal{X}$ that satisfy axioms 1–3 stated below. The generic element of Ω is given by ω. I denote the relation of weak preference with respect to $\omega \in \Omega$ by $x \succeq_\omega y$ and similarly for strict preference and indifference relations I use $\succ_\omega$ and $\sim_\omega$ respectively. Whenever I write $(x, y) \in \omega$ it denotes a relation of weak preference, that is $x \succeq_\omega y$. Real preferences of the consumer are denoted by $\omega^* \in \Omega$.

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

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

Axiom 3. *(Limited Indifference) Let $\omega \in \Omega$. For any $x \in \mathcal{X}$ set $\{y \in \mathcal{X} : x \sim_\omega y\}$ has an empty interior.*

Axioms 1 and 2 are standard axioms of utility theory. As $\mathcal{X}$ is metrizable and compact, it is also second countable and as such the theorem of (Debreu 1964) states that preferences that satisfy those two axioms can be represented by a continuous utility function[2]. The only other axiom that I assume, namely axiom 3 implies that the indifference curves in all possible preferences are thin and the intended interpretation of this axiom is that the consumer is highly unlikely to be indifferent between two randomly chosen alternatives. This is only assumed in order to simplify the model, because it allows me to mostly ignore the possibility that the consumer is indifferent between x, y and only consider the case that either $x \succ_\omega y$ or $y \succ_\omega x$.

Incomplete preferences

By incomplete preference relation I consider any finite binary relation on $\mathcal{X}$ that is transitive and reflexive. I denote generic incomplete preference relation by $\bar{\omega}$. For any given set of pairs of alternatives $A = \{(x_1, y_1), \ldots, (x_n, y_n)\} \subset \mathcal{X} \times \mathcal{X}$ i denote by $\bar{\omega}_A$ the smallest (with respect to inclusion) incomplete preference relation such that $A \subset \bar{\omega_A}$.

[2]Utility functions (and everything else in this model) are ordinal. However, it is easy to extend this model to cardinal utility, i.e. for choice under risk.

Two main operations I use to work with incomplete preference relations are given by definitions 1 – 2.

Definition 1. *Let* $\omega \in \Omega$ *and* $A \subset \mathcal{X}$*. I say that an incomplete preference relation* $\bar{\omega}_{|A}$ *defined by*

$$(x, y) \in \bar{\omega}_{|A} \iff ((x, y) \in \omega \wedge x, y \in A)$$

is the restriction of ω *to* A*.*

Definition 2. *Let* $\bar{\omega}$ *be an incomplete preference relation. I say that a set*

$$[\bar{\omega}] = \{\omega \in \Omega : (x, y) \in \bar{\omega} \implies (x, y) \in \omega\}$$

is the set of extensions of $\bar{\omega}$*.*

Similarly, for a set of incomplete preference relations $B = \{\bar{\omega}_i : i \in I\}$, its extension is defined by $[B] = \bigcup_{i \in I} [\bar{\omega}_i]$.

Let $A_1 = \{(x, y)\}$, $A_2 = \{(y, x)\}$, $A_3 = \{(x, y), (y, x)\}$. Abusing notation a little, whenever context is clear I denote an incomplete preference relations $\bar{\omega}_{A_1}$, $\bar{\omega}_{A_2}$ and $\bar{\omega}_{A_3}$ by respectively $x \succ y$, $x \prec y$ and $x \sim y$, and the sets of incomplete preference relations $\{x \succ y, x \sim y\}$, $\{x \prec y, x \sim y\}$ by respectively $x \succeq y, x \preceq y$.[3] Those are the incomplete preference relations where the consumer only knows a relation between a single pair of elements, so I call that those incomplete preference relations as just relations.

Similarly for $\bar{\omega} = \{(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n)\}$, I sometimes use the notation that $\bar{\omega} = x_1 \succ x_2 \succ \cdots \succ x_n$.

Finally, I equip Ω with a topology $\mathcal{T}$ generated by the family of extensions of strict relations, meaning that $\mathcal{T}$ is the smallest topology which for all $x, y \in \mathcal{X}$ satisfies $[x \succ y] \in \mathcal{T}$.[4] From now on I only consider Ω as a topological space $(\Omega, \mathcal{T})$.

[3] To avoid confusion note, that in this notation relation symbols like $\succ$ have no subscript denoting some $\omega \in \Omega$.

[4] This topology can be obtained as a standard product topology in a way that is very natural for ordinal preferences. Consider for each $\omega \in \Omega$ the function $f_\omega : \mathcal{X}^2 \to \{-1, 0, 1\}$, such that $f_\omega(x, y) = 1$ iff $x \succ_\omega y$, $f_\omega(x, y) = 0$ iff $x \sim_\omega y$ and $f_\omega(x, y) = -1$ iff $x \prec_\omega y$. Equip the set $\{-1, 0, 1\}$ in the topology, such that open sets are $\emptyset, \{-1\}, \{1\}, \{-1, 1\}, \{-1, 0, 1\}$. With this definition $\omega \in \Omega$ is continuous if and only if f_ω is. Now equip the whole space of continuous functions from $\mathcal{X}^2$ into $\{-1, 0, 1\}$ in the standard product topology and embed Ω into this space using the identification $\omega \to f_\omega$. It is an easy exercise to see that in this way I get the same topology.

Topological constructions on the space of preference relations are nothing new in economics; one well-known example of such a construction is given in (Kannai 1970).

Knowledge

Let $\mathcal{K} \subset \mathcal{X}$ denote a finite subset of alternatives, with generic elements $k, l, m \in \mathcal{K}$. I interpret elements of $\mathcal{K}$ as the alternatives that the consumer has already experienced and assume, that the real preference ranking of those elements is known, meaning that for each pair $k, l \in \mathcal{K}$ the consumer knows what is the relation between those two alternatives with respect to ω^*.

Relations known by the consumer are given by an incomplete preference relation $\bar{\omega^*}_{|\mathcal{K}}$, meaning that it is the preference relation that describes what the consumer knows regarding their own taste.

Accordingly, I define $\Omega(\mathcal{K})$ to be a subset of possible preferences that are not excluded by what the consumer already knows, meaning that $\Omega(\mathcal{K}) = [\bar{\omega^*}_{|\mathcal{K}}]$, that is $\Omega(\mathcal{K})$ is a set of all possible preferences that extend $\bar{\omega^*}_{|\mathcal{K}}$.

3 Consumer beliefs

Incomplete preference relation $\bar{\omega^*}_{|\mathcal{K}}$ represents what the consumer has directly learned from consumption of the alternatives in $\mathcal{K}$. The consumer also learns from consumption indirectly, by forming probabilistic beliefs over the set of possible preferences that respond to what they learned directly. Those beliefs are represented in the model by a probability measure, denoted μ, and a sigma field, denoted σ, on the space of possible preferences Ω.

Definition 3. *System of beliefs of the consumer is a measure space* (Ω, σ, μ).

The existence and properties of systems of consumer beliefs are the central point of this article.[5] The intended interpretation of $\mu(A)$ for a measurable $A \subset \Omega$ is a probability of $\omega^* \in A$. Note, that μ restricted to extensions of single relations, for example $[x \succ y]$ allows for another (equivalent) interpretation, meaning that

[5]For a sigma field of Borel subsets of Ω existence can be shown using Hahn-Banach theorem (easy exercise left to the reader) but I provide a constructive proof later in this section for the specific case which is of interest to me.

$\mu([x \succ y])$ is the ex-ante probability that after experiencing both x and y the consumer finds that x is indeed preferred to y.

As for sigma field σ, I always assume it to be a sigma field of Borel subsets of Ω. Defining a measure on the σ-field of Borel sets is a typical way of ensuring that the measure is in some sense compatible with the topology of the underlying space. Proposition 1 below shows two important facts regarding this sigma field, namely it establishes the measurability of the necessary subsets of Ω, and gives the set that generates this sigma field.

Proposition 1. *Let σ_B denote a Borel sigma field on Ω.*

1. *Let $C = \{[x \succ y] : x, y \in \mathcal{X}\}$. Then $\sigma_B = \sigma(C)$, that is the smallest sigma field containing all sets in C.*

2. *Let $B = \{\bar{\omega}_i, i \in \mathbb{N}\}$ be a set of incomplete preference relations. Then $[B] \in \sigma_B$.*

Proof. The family of all finite intersections of sets in C is a base of the topology on Ω. Moreover, as $\mathcal{X}$ is second countable, it is separable and as such it has a countable dense subset. Let $A \subset \mathcal{X}$ be this subset. By axiom 2 each $\omega \in \Omega$ is continuous and as such is uniquely determined by its relations on $A \times A$. Therefore the family of all finite conjunctions of conditions on $A \times A$ also is a base of the topology on Ω. Now it follows that a sigma field generated by the family of sets $\{[x \succ y] : x, y \in \mathcal{X}\}$ is equal to σ_B. The second point of the proposition follows from the first one.

□

By definition 3, system of beliefs of the consumer is not dependent on $\mathcal{K}$. This is because system of beliefs does not intend to represent the current beliefs of the consumer, but rather how those beliefs respond to any knowledge that the consumer might have. The current beliefs of the consumer are defined in definition 5.

Definition 4. *Let μ and some measurable $A \subset \Omega$ be given and denote*

$$\bar{K} = \{k \succ l : k \succ l \subset \bar{\omega^*}_{|\mathcal{K}}\}.$$

I define

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

Definition 5. *System of conditional beliefs of the consumer is a measure space* $(\Omega(\mathcal{K}), \sigma_{\mathcal{K}}, \mu_{\mathcal{K}})$*, where* $\sigma_{\mathcal{K}} = \{A \cap \Omega(\mathcal{K}) : A \in \sigma\}$.

Definition 4 is just a normal definition of a conditional probability, that is $\mu_{\mathcal{K}}(A)$ is a probability that $\omega^* \in A$ conditionally on the knowledge that $\omega^* \in \Omega(\mathcal{K})$. The only non-standard element in the definition 4 is that in case there are some $k, l \in \mathcal{K}$ such that the consumer is revealed to be indifferent between k and l, meaning that $k \sim l \subset \bar{\omega^*}_{|\mathcal{K}}$, this relation is ignored when calculating the conditional probability. It is because as I show later on, for any $x, y \in \mathcal{X}$ set $[x \sim y]$ is of measure zero. Definition 4 handles this case by specifying that the consumer does not learn anything about alternatives other than x, y from learning that $x \sim y$. In other words in this case there is no indirect learning, but direct learning still takes place because the domain is restricted to those $\omega \in \Omega$ for which $x \sim_\omega y$ holds.

Axioms 4–6 below state some primitive assumptions regarding $\mu_{\mathcal{K}}$ that from now on I always assume to be satisfied. These axioms also hold for μ because $\mu = \mu_\emptyset$.

Axiom 4. *(Non-degeneracy) Let* $U \subset \Omega$ *be open and nonempty. Then* $\mu_{\mathcal{K}}(U) > 0$.

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

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

Axiom 4 states, that for any pair $x, y \in \mathcal{X}$ with $x \neq y$ it is ex ante possible that either x is preferred to y or vice versa. In other words, I assume that as long as at least one of the alternatives has not yet been consumed, the consumer is never totally certain which one they prefer.

In general, these probabilities are not equal because of indirect learning from consumption. Indirect learning is incorporated into the model by axiom 5. It is just a typical continuity axiom, that states that $\mu_{\mathcal{K}}$ restricted to sets like $[x \succ y]$

and treated as a function of x, y is continuous. Therefore, by consumption of any $x \in \mathcal{X}$ the consumer also learns that if x is preferred to y then any alternative x' sufficiently similar to x should also be preferred to y.

The restriction of axiom 5 to the pairwise non-equal elements is only necessary, because by axiom 6 $\mu([x \sim y]) = 0$ for $x \neq y$, however it is clear that $\mu([x \sim x]) = 1$. So for the case $x = y$ the measure is inherently discontinuous. In order to avoid the necessity of special treatment of the sets like $[x \sim x]$ everywhere, from now on I use the convention that $\mu([x \succ x]) = \frac{1}{2}$ and $\mu([x \sim x]) = 0$. With such convention, axiom 5 holds for any x, y, z.[6]

Axiom 6 is only introduced to allow me to ignore indifference relations, which greatly simplifies proofs and notation. This axiom is clearly connected to axiom 3, however it is not implied by it. Axiom 3 restricts the space Ω, whereas axiom 6 restricts measures on Ω.

I am now ready to state the main result of this section, which is given in theorem 1. Proof of this theorem is in the appendix.

Theorem 1. *For all $n \in \mathbb{N}_+$ and $A = \{(x_1, y_1), \dots (x_n, y_n) : x_i, y_i \in \mathcal{X}\}$, let $A_1 = A \cup \{(x_{n+1}, y_{n+1})\}, A_2 = A \cup \{(y_{n+1}, x_{n+1})\}$ and assume the values of set function $\mu_0([\bar{\omega}_A]) > 0$ to be given such that for all x_{n+1}, y_{n+1} those values satisfy*

$$\mu_0(\bar{\omega}_{A_1}]) + \mu_0([\bar{\omega}_{A_2}]) = \mu_0([\bar{\omega}_A]),$$

and for all $x, y \in \mathcal{X}$

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

Then there exists a unique probabilistic measure μ defined on the whole σ-field of Borel subsets of Ω such that $\mu([\bar{\omega}_A]) = \mu_0([\bar{\omega}_A])$ for all incomplete preferences $\bar{\omega}_A$. Moreover, μ satisfies axioms 4-6 if and only if μ_0 also satisfy those axioms.

Theorem 1 is a measure definition theorem. It is analogous to the well known Caratheodory extension theorem and it states that in order to identify the system of consumer beliefs, it is sufficient to define μ only on the extensions sets of finite incomplete preference relations. Moreover, the conditions of the theorem are minimal. The only requirement is for the probabilities to sum up to one.

[6] Of course $\mu([x \sim x]) = 1$. Formally, this convention means that I define a set function $\tilde{\mu}$ as $\tilde{\mu}([x \sim x]) = 0$, $\tilde{\mu}([x \succ x]) = \frac{1}{2}$ and $\tilde{\mu}(A) = \mu(A)$ otherwise, and that unless specified otherwise, whenever I write μ, in reality I refer to $\tilde{\mu}$.

This result is also empirically significant, due to the fact that $\mu([x \succ y])$ can be interpreted as the probability that after consumption it will be revealed that x is preferred to y and similarly, for $\bar{\omega} = \{(x, y), (y, z)\}$ value $\mu([\bar{\omega}])$ is the probability that after the consumption it will be revealed that both x is preferred to y and y to z. As such, those sets are easily interpretable and understandable.

The following corollary 1 provide an equivalent method of defining μ.

Corollary 1. *Let for all $x, y \in \mathcal{X}$ and all finite $\mathcal{K} \subset \mathcal{X}$ the values of $\mu^0_{\mathcal{K}}([x \succ y]) > 0$ to be given and satisfy*

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

There exists a unique probabilistic measure μ defined on the whole Borel σ-field, such that $\mu([x \succ y]) = \mu^0_{\emptyset}([x \succ y])$ and $\mu_{\mathcal{K}}([x \succ y]) = \mu^0_{\mathcal{K}}([x \succ y])$ for all $x, y \in \mathcal{X}$.

Proof. For any $x, y \in \mathcal{X}$ note that $\mu([\bar{\omega^*}_{|\mathcal{K}}] \cap [x \succ y]) = \mu_{\mathcal{K}}([x \succ y])\mu(\Omega(\mathcal{K}))$. Now for it to follow from theorem 1 I just need to show that I am able to calculate $\mu(\Omega(\mathcal{K}))$ using values of $\mu_{\mathcal{K}}([x \succ y])$ only. I show it by induction on number of elements in $\mathcal{K}$. Let $|\mathcal{K}| = 2$. Then $\mu(\Omega(\mathcal{K})) = \mu_{\emptyset}([x_1 \succ x_2])$ for some $x_1, x_2 \in \mathcal{X}$. Now assume I am given $\mu(\Omega(\mathcal{K}))$ for $|\mathcal{K}| = n$ and let $|\mathcal{K}'| = n + 1$. Moreover, let $\bar{\omega^*}_{|\mathcal{K}} = x_1 \succ \cdots \succ x_n$ and $\bar{\omega^*}_{|\mathcal{K}} = x_1 \succ \cdots \succ x_j \succ x_{n+1} \succ x_{j+1} \succ \cdots \succ x_n$.

From definition 4, I have $\mu(\Omega(\mathcal{K}')) = \mu(\Omega(\mathcal{K}))\mu_{\mathcal{K}}([x_j \succ x_{n+1} \succ x_{j+1}])$, where $\mu_{\mathcal{K}}([x_j \succ x_{n+1} \succ x_{j+1}]) = 1 - \mu_{\mathcal{K}}([x_{n+1} \succ x_j] \cup [x_{j+1} \succ x_{n+1}])$. Given that $x_j \succ x_{j+1} \subset \bar{\omega^*}_{|\mathcal{K}}$, the sets $[x_{n+1} \succ x_j], [x_{j+1} \succ x_{n+1}]$ are disjoint in $\Omega(\mathcal{K})$. Therefore $\mu_{\mathcal{K}}([x_j \succ x_{n+1} \succ x_{j+1}]) = 1 - \mu_{\mathcal{K}}([x_{n+1} \succ x_j]) - \mu_{\mathcal{K}}([x_{j+1} \succ x_{n+1}])$.

□

Theorem 1 together with corollary 1 establishes the model for the identification of the system of beliefs of the consumer. However, the question remains how general is this model, meaning which systems of beliefs can be supported by this model. In order to answer this question, I consider an expected preference relation, which is given in definition 6.

Definition 6. *Let $\mu_{\mathcal{K}}$ be given. The relation $\omega_{\mathcal{K}}$ defined by $(x, y) \in \omega_{\mathcal{K}} \iff \mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}$ is the expected preference relation.*

Expected preference relation given in definition 6 is defined such that x is weakly preferred to y with respect to this relation if and only if the consumer expects that the it will prove to be the case after the consumption. Despite the name, it does not have to be the case that $\omega_{\mathcal{K}}$ is a preference relation because it is not instantly clear whether it is transitive. I discuss the transitivity of $\omega_{\mathcal{K}}$ in further detail and provide the conditions for the transitivity to hold in section 4.

For a given $\omega \in \Omega(\mathcal{K})$ I say that $\mu_{\mathcal{K}}$ represents ω if and only if $x \succ_{\omega} y \iff \mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}$. Theorem 2 states, that for each $\omega \in \Omega(\mathcal{K})$ it is possible to find $\mu_{\mathcal{K}}$ representing it. Proof is provided in the appendix.

Theorem 2. *Let $\mathcal{K}$ be fixed. For all $\omega \in \Omega(\mathcal{K})$ there exists a probability measure $\mu_{\mathcal{K}}$ satisfying axioms 4–6 such that $\mu_{\mathcal{K}}$ represents ω.*

Firstly, theorem 2 finally provides the proof that probabilistic measures satisfying axioms 4–6 exist. Moreover, theorem 2 shows that the model of consumer beliefs is very general, as every possible preference relation can be represented in this way. This representation is obviously not unique, therefore two different measures can represent the same expected preference relation, but at the same time for a different preference relation those two measures might differ.

4 Preference reversal

Preference reversal (for example Lichtenstein and Slovic 1971, Stalmeier, Wakker, and Bezembinder 1997) is a paradox, which occurs in experiments during which the subjects have two tasks. Firstly, in the choice task the subjects are asked to choose which alternative out of x, y do they prefer. Secondly, in the valuation tasks, they are asked to state the maximum amount they would be willing to pay for x and y. A typical pattern of behaviour is, that the subjects choose x in the choice task, but assign a higher value to y in valuation task.

It is an important paradox for economic theory, since it seems to imply that there is no optimizing behaviour whatsoever behind consumer choices. Empirical studies seem to suggest that there is a link between preference reversal and preference discovery. However, the nature of this link is uncertain, as it is unclear why taste uncertainty should lead to preference reversals when no new taste information is acquired in between. Indeed, no contemporary model of taste uncertainty predicts preference reversal. It is the case because preference reversal implies that either there is a different choice procedure applied in choice and valuation tasks, or that the preferences under taste uncertainty are not transitive.

In this section I provide equivalent conditions for two hypotheses regarding the reason for the link between preference discovery and preference reversal. The first hypothesis is, that there is the same choice procedure in both tasks, which I assume to be given by the expected preference relation, but that the consumer preferences under taste uncertainty are intransitive. The second hypothesis is, that there is a different procedure between those two tasks. More precisely, I assume that in the choice task, the consumer compares the alternatives directly, using the expected preference relation. However in the valuation task, the consumer compares the choices indirectly, by comparison with an unrelated alternative. This indirect preference relation is given by definition 7.

Definition 7. *Let $z \in \mathcal{X}$. The indirect preference relation ω_z with respect to reference point z is defined as $x \succeq_{\omega_z} y \iff \mu_{\mathcal{K}}([x \succeq z]) \geq \mu_{\mathcal{K}}([y \succeq z])$.*

Indirect preference relation compares both alternatives x and y to some unrelated alternative z, which I treat as a reference point, and states that x is indirectly preferred to y if it is more likely than y to be better than z. Note, that for all reference points $z \in \mathcal{X}$ indirect preferences are clearly transitive and utility function $u_z(x) = \mu_{\mathcal{K}}([x \succ z])$ represents ω_z. I call this utility function an indirect utility.

Two key properties of conditional measure $\mu_{\mathcal{K}}$, namely coherence and weak coherence, are given in definitions 8 and 9 below.

Definition 8. *$\mu_{\mathcal{K}}$ is coherent if $(\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}) \implies (\forall_{z \in \mathcal{X}} : \mu_{\mathcal{K}}([x \succeq z \succeq y]) \geq \mu_{\mathcal{K}}([y \succeq z \succeq x]))$.*

Definition 9. *For $x, y \in \mathcal{X}$ let $A^x_y = \{z \in \mathcal{X} : \mu([z \succeq x]) \geq \frac{1}{2} \vee \mu([y \succeq z]) \geq \frac{1}{2}\}$. $\mu_{\mathcal{K}}$ is weakly coherent if $(\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}) \implies (\forall_{z \notin A^x_y} : \mu_{\mathcal{K}}([x \succeq z \succeq y]) \geq \mu_{\mathcal{K}}([y \succeq z \succeq x]))$.*

Both coherence and weak coherence are properties that restrict the correlations between different beliefs of the consumer. Coherence demands, that if the consumer believes that it is more likely that x would after consumption prove to be preferred to y, then for any unrelated z it also should be more likely that it would prove that the revealed ordering of x, y, z will be $x \succ_{\omega^*} z \succ_{\omega^*} y$ then $y \succ_{\omega^*} z \succ_{\omega^*} x$. The only difference between those two properties is that weak coherence restrict the domain to which this restriction applies to those alternatives z which are believed to be between x and y in the preference ranking.

Theorem 3 gives an answer to the first hypothesis. It shows that weak coherence is the condition equivalent to transitivity of the expected preference relation. Moreover, theorem 3 shows that weakly coherent measures exist, and can be found to represent any possible $\omega \in \Omega(\mathcal{K})$. Proof of this theorem is in the appendix.

Theorem 3. *The following hold.*

1. *Expected preference $\omega_{\mathcal{K}}$ is transitive if and only if $\mu_{\mathcal{K}}$ is weakly coherent.*
2. *For each $\omega \in \Omega(\mathcal{K})$ there exists a weakly coherent conditional measure $\mu_{\mathcal{K}}$ such that $\mu_{\mathcal{K}}$ represents ω.*
3. *There exist μ such that for all $\mathcal{K}$ conditional measure $\mu_{\mathcal{K}}$ is weakly coherent.*

Theorem 3 shows that expected preference relation is transitive if and only if it is weakly coherent. Therefore we can only expect to see preference reversals as a result of a lack of transitivity of the expected preferences if $\mu_{\mathcal{K}}$ is not weakly coherent.

Now the question remains what is the equivalent condition for indirect preference relations to coincide for all reference points, and for it to be equivalent to expected preference relation. This condition is provided by proposition 2.

Proposition 2. *Let $z \in \mathcal{X}$ and $\mathcal{K}$ be fixed. Then*

1. $\forall_{z' \in \mathcal{X}} : \omega_z = \omega_{z'}$ *if and only if $\mu_{\mathcal{K}}$ is coherent.*
2. $\forall_{z \in \mathcal{X}} : \omega_z = \omega_{\mathcal{K}}$ *if and only if $\mu_{\mathcal{K}}$ is coherent.*

Proof. Assume $\mu_{\mathcal{K}}$ is coherent and without loss of generality fix $x, y \in \mathcal{X}$ such that $\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}$. First note, that coherence is equivalent to the condition, that $\mu_{\mathcal{K}}([x \succ y]) \geq \frac{1}{2} \implies \forall_{z \in \mathcal{B}} \mu_{\mathcal{K}}([x \succ z]) \geq \mu_{\mathcal{K}}([y \succ z])$. Indeed

$$\mu_{\mathcal{K}}([x \succ z]) \geq \mu_{\mathcal{K}}([y \succ z]) \iff \mu_{\mathcal{K}}([y \succ x \succ z]) + \mu_{\mathcal{K}}([x \succ y \succ z]) + \mu_{\mathcal{K}}([x \succ z \succ y]) \geq$$

$$\geq \mu_{\mathcal{K}}([x \succ y \succ z]) + \mu_{\mathcal{K}}([y \succ x \succ z]) + \mu_{\mathcal{K}}([y \succ z \succ x]) \iff$$

$$\iff \mu_{\mathcal{K}}([x \succ z \succ y]) \geq \mu_{\mathcal{K}}([y \succ z \succ x]).$$

Therefore coherence is equivalent to the fact, that for any $z_1, z_2 \in \mathcal{X}$ I 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 $\forall_{z_1,z_2\in\mathcal{B}} : \omega_{z_1} = \omega_{z_2}$. I will show that $\mu_{\mathcal{K}}$ is coherent. Again fix $x, y \in \mathcal{X}$ such that $\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}$. Note that for any $z \in \mathcal{X}$, u_z and u_y represent the same preferences. Since $u_y(x) \geq u_y(y)$, therefore $u_z(x) \geq u_z(y)$, and therefore coherence holds.

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

□

Proposition 2 gives an answers to the second hypothesis. It states, that coherence is the equivalent condition for expected preference relation to give the same preference ranking as indirect preference relations. Another consequence of this proposition is given by corollary 2.

Corollary 2. *Assume $\mu_{\mathcal{K}}$ is weakly coherent. Then for every pair $z_1, z_2 \in \mathcal{X}$ such that $z_1 \sim_{\mathcal{K}} z_2$ I have $\omega_{z_1} = \omega_{z_2}$.*

Proof. Note, that for the case $z_1 \sim_{\mathcal{K}} z_2$ definitions 8 and 9 coincide, due to transitivity of $\omega_{\mathcal{K}}$ shown in theorem 3. Therefore by proposition 2 weak coherence of $\mu_{\mathcal{K}}$ implies that for any pair $z_1 \sim_{\mathcal{K}} z_2$ preference relations $\omega_{z_1}, \omega_{z_2}$ coincide.

□

Corollary 2 shows that for weakly coherent conditional measure, the indirect preference relation for z, z' that are expected to be just as good as one another coincide. As such, for a weakly coherent $\mu_{\mathcal{K}}$ it is possible to interpret indirect preferences as solving the problem of maximization of the probability of obtaining at least the level of utility provided by z.

Proposition 2 states that cherence is necessary for all indirect preferences to coincide and be equal to expected preference relation. However coherence is a very demanding condition as shown by the following proposition 3.

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

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

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

□

Proposition 3 shows that coherent measures do not exist. Therefore, we should expect that direct and indirect choice might always lead to different rankings of alternatives.

5 Applications and extensions

Value of information

My construction is restrictive from the perspective of the study of learning. Firstly, I assume that consumption perfectly reveals preference rankings between alternatives, so direct learning from consumption is trivial. Secondly, the construction provided by theorem 1 does not restrict indirect learning in any way beside the demand for continuity of μ, and as shown by corollary 1, values of $\mu_{\mathcal{K}}$ and $\mu_{\mathcal{K}\cup\{x\}}$ are not implicitly connected. Finally, definition 4 implies that the consumer perfectly anticipates changes in $\mu_{\mathcal{K}}$ that result from a new information. As a result, to study dynamic properties of the model such as learning or experimentation, additional structure is required.

That said, we can characterize experimental preferences of the consumer by how much additional information the consumer expects to obtain from consumption of a given alternative. This concept is formalized by definition 10.

Definition 10. *Let $x, y \in \mathcal{X}$ and assume $\mathcal{K}$ is given. Denote $\mathcal{K}_x = \mathcal{K} \cup \{x\}$ and $\mathcal{K}_y = \mathcal{K} \cup \{y\}$. I say that x is experimentally weakly preferred to y, denoted by $x \succeq_{\omega_E} y$ if and only if $E_{\mu_{\mathcal{K}}}[\mu(\Omega(\mathcal{K}_x))] \leq E_{\mu_{\mathcal{K}}}[\mu(\Omega(\mathcal{K}_y))]$.*

Note, that $\mu(\Omega(\mathcal{K}))$ is a natural measure of the taste uncertainty that is yet to be resolved, and consumption of x resolves at least as much uncertainty then y if $\mu(\Omega(\mathcal{K}_x)) \leq \mu(\Omega(\mathcal{K}_y))$. However, both of those values are ex ante unknown, as they depend on what the revealed relations between x, y and the elements of $\mathcal{K}$ will turn out to be after the consumption. Therefore ex ante, $\mu(\Omega(\mathcal{K}_x))$ and $\mu(\Omega(\mathcal{K}_y))$ are both random variables and definition 10 states that x is experimentally weakly preferred to y if the expected value of the random variable $\mu(\Omega(\mathcal{K}_x))$ is not higher than that of $\mu(\Omega(\mathcal{K}_y))$, meaning that the consumer expects consumption of x to resolve at least as much uncertainty as of y.

The following proposition 4 provides a representation for ω_E.

Proposition 4. *Experimental preferences ω_E are always complete, transitive, continuous and reflexive. Moreover let $\mathcal{K} = \{x_1, \ldots, x_n\}$ be such that $x_i \succeq_{\omega^*} x_{i+1}$ for all i. Then the utility function $u_E(x) = 1 - \sum_{i=1}^{n-1} \mu^2_{\mathcal{K}}([x_i \succ x \succ x_{i+1}]) - \mu^2_{\mathcal{K}}([x \succ x_1]) - \mu^2_{\mathcal{K}\sqcup}([x_n \succ x])$ represents ω_E.*

Proof. Since u_E as defined in the statement of the theorem is continuous, the second part of the theorem implies the first part. Therefore I only need to prove that u_E represents ω_E. From definition 4 I have the following

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

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

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

Therefore

$$x \succeq_{\omega_E} y \iff (1 - u_E(x))\mu(\Omega(\mathcal{K})) \leq (1 - u_E(y))\mu(\Omega(\mathcal{K})) \iff u_E(x) \geq u_E(y).$$

□

Proposition 4 gives a very natural utility function for experimental preferences. Lets denote the probability of x being in i-th position in the ranking of known alternatives as p_i, meaning that $p_i = \mu_{\mathcal{K}}([x_{i-1} \succ x \succ x_{i+1}])$ for $i = 2, \ldots, n$, with $p_1 = \mu_{\mathcal{K}}([x \succ x_1])$ and $p_{n+1} = \mu_{\mathcal{K}}([x_n \succ x])$. Then $1 - u_E$ is simply a quadratic form $\sum_{i=1}^{n+1} p_i^2$ and for example the maximal element with respect to ω_E is the one such that $p_i = \frac{1}{n+1}$ for all i (if such an element exists), meaning that each position in the resulting preference ranking is equally probable.

Other measurable sets

In application to preference reversal, I am mostly interested in probabilities for extension sets of some incomplete preference relations, especially those defined over a single pair of elements like $\mu_{\mathcal{K}}([x \succ y])$. It is because of the interpretation for those sets as a probability that after the consumption of both x, y the revealed relation between those elements it will prove to be the case that x is preferred to y.

However, the Borel sigma field contains many other interesting subsets of Ω. To fix the attention, lets consider the case of $\mathcal{X} = [0,1]^2$ with euclidean metric. The elements of $\mathcal{X}$ might represent different movies with each dimension being an attribute of the movie, for example the first axis might be how serious the movie is, so that the movie is a comedy for values of attribute close to 0 and a documentary close to 1, and the second one how packed with action it is, meaning that a documentary would be close to 0 on this axis, whereas a fast-paced action thriller would be close to 1.

Now, from a perspective of a streaming platform, we might be interested in how the consumer perceives their preference for action movies, meaning that we could be interested in the probability that consumer preferences are monotone with respect to the second dimension.

Definition 11. *Let $\mathcal{X} = [0,1]^2$. Preference relation $\omega \in \Omega$ is monotone with respect to the second dimension if and only if $\forall_{x,y,z \in [0,1]}(x \geq y) \implies ((z,x) \succeq_\omega (z,y))$.*

I define a function $m : \Omega \to \{0,1\}$ such that $m(\omega) = 1$ if ω satisfies definition 11 and $m(\omega) = 0$ otherwise.

Proposition 5. *Let σ_B be a Borel sigma field on Ω. Then*

$$\{\omega \in \Omega : m(\omega) = 1\} \in \sigma_B.$$

Proof. Let D be a countable, dense subset of $[0,1]$. Because all $\omega \in \Omega$ are continuous, it is enough to check definition 11 on $D \times D$.

It is obvious, that

$$\{\omega \in \Omega : m(\omega) = 1\} = \bigcap_{d_1 \in D} \bigcap_{d_2 \in D} [(d_1, 1) \succ (d_1, d_2)].$$

Because sigma fields are closed with respect to countable intersections, proof is finished.

□

Proposition 5 shows, that the subset of preferences monotone in one direction is measurable, therefore there exists a probability that the real preference relation of the consumer is monotone.

Another example we might be interested in is the comparison between different genres of movies. For example, let $A, C \subset [0,1]^2$ be the subsets of respectively action movies and romantic comedies. We might be interested in whether the consumer prefers one category over another, as defined in definition 12

Definition 12. *Let $A, C \subset \mathcal{X}$ and $\omega \in \Omega$ be given.*

- *Set A is strictly preferred to set C if $\forall_{a \in A} \forall_{c \in C}\ a \succeq_\omega c$.*
- *Set A is weakly preferred to set C if $\forall_{c \in C} \exists_{a \in A}\ a \succeq_\omega c$.*

I again define functions $s, w : \Omega \to \{0, 1\}$ such that $s(\omega) = 1, w(\omega) = 1$ if A is respectively strictly and weakly preferred to C with respect to ω, and 0 otherwise. Again, subsets of preferences for which action movies either strictly or weakly dominate romantic comedies are both measurable.

Proposition 6. *Let σ_B be a Borel sigma field on Ω. Then*

$$\{\omega \in \Omega : w(\omega) = 1\} \in \sigma_B, \quad \{\omega \in \Omega : s(\omega) = 1\} \in \sigma_B.$$

Proof. Let D be a countable dense subset of $[0,1]^2$ and denote $D_A = D \cap A$ and $D_C = D \cap C$. Note that

$$\{\omega \in \Omega : s(\omega) = 1\} = \bigcap_{c \in D_A} \bigcap_{a \in D_A} [a \succ c],$$

which is an element of σ_B because it is closed with respect to countable intersections. Similarly

$$\{\omega \in \Omega : w(\omega) = 1\} = \bigcap_{c \in D_C} \bigcup_{a \in D_A} [a \succ c],$$

and the proof is finished.

□

It is not possible to give a full characterization of measurable subsets of Ω, but there are many more interesting measurable subsets of Ω. Finally, note that by definition 5 the results in both proposition 5 and 6 also hold for conditional probability measures for all possible $\mathcal{K}$.

Extended model of consumer knowledge

Consumer knowledge in section 2 is defined as a subset of alternatives, $\mathcal{K}$ for which the consumer knows real preference relation. This definition is sufficient for my analysis of preference reversal, but my model allows for a way more general definition of consumer knowledge.

A function $b : \Omega \to \{0, 1\}$ is a predicate on Ω. For any predicate[7], I define a set of those elements of Ω which satisfy the predicate as $[b] = \{\omega \in \Omega : b = 1\}$ and knowledge of the consumer as a finite set of binary predicates $\mathcal{K} = \{b_1, \ldots, b_n\}$. The subset of possible preferences which agree with consumer knowledge $\Omega(\mathcal{K})$ can be now defined as

$$\Omega(\mathcal{K}) = \left\{ \omega \in \Omega : \omega \in \bigcap_{i=1}^{n} [b_i] \right\}.$$

It is an easy exercise to extend all the other definitions in the model to this formulation of consumer knowledge. This is a strictly more general formulation. Indeed, assume that set $\mathcal{K}' = \{k_1, \ldots, k_n\}$ is a set of known alternatives. We can easily write a predicate

$$b_{ij}(\omega) = \begin{cases} 1, & \text{if } k_i \succeq_\omega k_j \iff k_i \succeq_{\omega^*} k_j \\ 0, & \text{otherwise.} \end{cases}$$

Now it suffices to define $\mathcal{K} = \{b_{ij} : i, j \leq n\}$ to obtain the same knowledge representation as in section 2, meaning that for example $\Omega(\mathcal{K})$ is the subset of preferences which agree with ω^* on the subset of known alternatives. However, this formulation can easily incorporate more general knowledge that the consumer might have. For example if $\mathcal{X} = [0, 1]^n$, then in some situations the consumer

[7]Clearly, not all predicates make sense as something that the consumer could know. It is possible to restrict the domain of possible predicates by defining a first order logic structure with universe $\mathcal{X}$ such that set of models for this structure is Ω and allowed predicates are those obtained from valuation functions of well-formed formulas in this structure. I refrain from doing this, and instead quietly assume that all the predicates are reasonable.

might know that their preferences are monotone with respect to some dimension. I have defined a predicate m representing monotone preferences in the previous section. Proposition 5 shows that $[m]$ is measurable, and as such we can include it in $\mathcal{K}$.

Adding a predicate to all of the knowledge sets of the consumer is mathematically equivalent to introducing an additional axiom to the model. Therefore, this general knowledge representation also allows us to easily introduce additional axioms on possible preferences to the model. For example, adding monotonicity as defined in definition 11 to the model is mathematically equivalent to adding a predicate m to each possible knowledge set $\mathcal{K}$ of the consumer, i.e. treating it as $\mathcal{K} \cup \{m\}$.

6 Discussion

Contemporary literature on taste uncertainty mostly follow (Kreps 1979) and use preference for flexibility in order to obtain a subjective state space that represents how the taste uncertain consumer perceives own preferences. Both (Piermont, Takeoka, and Teper 2016) and (Cooke 2017) provide extensions of the model of (Kreps 1979), which conditions the resolution of the taste uncertainty on consumption. There are a technical differences between those two models, for example (Piermont, Takeoka, and Teper 2016) considers ordinal preferences defined over infinite horizon choice problems whereas in model of (Cooke 2017) preferences are cardinal and the objects of choice are just a single period menu and consumption pairs.

There is a number of technical differences between by model and both of those contributions and my model. For example, like (Cooke 2017) I only consider a single period consumption, but I consider ordinal preferences like (Piermont, Takeoka, and Teper 2016). In contrast to both of these models, I do not consider preferences over menus and do not assume preference for flexibility. However, the biggest difference is conceptual. Both (Piermont, Takeoka, and Teper 2016) and (Cooke 2017) consider the consumer as equipped with some interim preferences, which are then represented using a subjective probability measure over some subjective state space. I reverse this process and consider consumer beliefs, represented by a probability measure, to be a primitive element of the model which can be used to define different preference relations. Finally, probability measures which I obtain are much more general.

Conceptually most similar to what I do is the case-based decision theory of (Gilboa and Schmeidler 1995). I also consider the consumer that evaluates available alternatives based on the consequences of their past choices, and indirect learning in my model also works by evaluating current choices by similarity to known alternatives. However, I apply these ideas to the beliefs of the consumer, not to the choice itself.

I am also methodologically close to theories of preference construction, of which (Lichtenstein and Slovic 2006) provide a comprehensive summary. I also consider the consumer who evaluates the alternatives using different choice procedures, the choice of which might depend on how the question is asked or context. However, there are two key differences between my model and the literature on preference construction. Firstly, I consider different choice procedures only when the alternatives are unknown to the consumer, meaning that those procedures should coincide for the goods which are already known or under perfect information. Secondly, I assume there exists some real preference relation which is discovered by consumption. Therefore in my model, consumer preferences are not history or context dependent, even if the choices during preference discovery might be.

The literature on incomplete preferences, (e.g. Bewley 2002, Ok, Ortoleva, and Riella 2012, Huang et al. 2014) as well as on random utility, are also closely related to my model. Indeed the information that is available to the consumer in my model takes the form of an incomplete preference relation. Especially notable from my perspective is the contribution of (Huang et al. 2014), as it applies the notion of an evidence distance that is conceptually very close to how I construct the probability measures. The conditional probability measure that is the central object of my model can be understood as a prior beliefs of the consumer over the possible resolution of the incompleteness. When it comes to random utility, the data in my model is similar to their, meaning that the appropriate question to ask is what is the probability that one alternative is better than another, and not just which alternative is better. However, the interpretation of this probability is different. In random utility, this probability represents the frequency of the alternative being chosen over another one, whereas in my interpretation, this is just an ex-ante strength of belief, that after the consumption of both alternatives, this one would be revealed to be better.

Dominant explanations of preference reversal in the literature are those of firstly (Tversky, Sattath, and Slovic 1988) who consider it to be a result of different choice procedures between tasks, and secondly of (Sugden 2003) who explains it as a result of a small changes of reference point between tasks. My analysis is not inconsistent with any of those explanations, but rather complementary to both. Indeed for beliefs of the consumer which satisfy weak coherence, my analysis shows that preference reversal might occur both as a result of different procedures, because direct and indirect choice do not coincide, as well as different reference points between tasks, because indirect choice is reference dependent. My analysis just answers the question as to why preferences of the consumer should be either procedure or reference point dependent only in the case when the consumer is suffering from taste uncertainty.

My analysis of preference reversal is also consistent with the findings of (Bostic, Herrnstein, and Luce 1990), who has shown that the ratio of preference reversals decrease substantially when instead of asking the subjects to give a valuation of of both alternatives, the valuations are elicited by a series of questions comparing a given alternative to a given amount. In this way, subjects use direct comparisons, rather than indirect, in the valuation task. This experiment suggests that preference reversals may indeed be a result of the discrepancy between direct and indirect comparisons.

To my knowledge, the only other theoretical model which considers the connection between taste uncertainty and paradoxes of choice is (G. Loomes, Orr, and Sugden 2009), where the authors construct a reference dependent prospect theory model with taste uncertainty, and consider what is the impact of taste uncertainty on the WTP/WTA disparity. However, my model treats taste uncertainty in a much more general way, is less restricted and does not use either prospect theory or, outside of indirect comparisons, reference dependence.

A Proof of theorem 1

A.1 Supporting facts and definitions

For a given incomplete preference relation $\bar{\omega} = \{(x_1, y_1), \dots (x_n, y_n)\}$, I define its length as $l(\bar{\omega}) = n$ and a set $\text{cp}(\bar{\omega}) = \{(x_1, y_1), \dots (x_n, y_n)\}$ of the pairs of points on which the relation is defined.

Definitions 13–15 introduce the concept of representation and its two properties. I mainly work with extension sets of incomplete preference relations, but representations allow me to study more sets using the same techniques.

Definition 13. *Fix $U \subset \Omega$ and let $R = \{\bar{\omega}_i : i \in I\}$ for arbitrary $I \subset \mathbb{N}$ and incomplete preferences $\bar{\omega}_i$. R is a representation of U if $U = \bigcup_{\bar{\omega}_i \in R} [\bar{\omega}_i]$.*

Definition 14. *Representation $R = \{\bar{\omega}_i : i \in I\}$ is disjoint if $i_1 \neq i_2$ implies that $[\bar{\omega}_{i_1}] \cap [\bar{\omega}_{i_2}] = \emptyset$.*

Definition 15. *Let R_1, R_2 be two representations of some set $U \subset \Omega$. R_1 is subordinate to R_2 if for all $\bar{\omega} \in R_1$ there is $\bar{\omega}' \in R_2$ such that $[\bar{\omega}] \subset [\bar{\omega}']$.*

It is clear that the representations are not unique, and that not every set has a representation because I only consider incomplete preferences defined over finite subsets of alternatives. Definitions 16 and 17 give the two main operations which I use to work with representations.

Definition 16. *Let a finite subset $A \subset \mathcal{X} \times \mathcal{X}$ and a finite disjoint representation R of $U \subset \Omega$ be given. Moreover, for all $\bar{\omega} \in R$ denote $A_{\bar{\omega}} = A \cup cp(\bar{\omega})$. A set*

$$\tilde{R} = \bigcup_{\bar{\omega} \in R} \{\bar{\omega}'_{|A_{\bar{\omega}}} : \bar{\omega}' \in [\bar{\omega}]\}$$

is a partition of R with respect to A.

Definition 17. *Let incomplete preference relation $\bar{\omega}$ be given and let R be a finite disjoint representation of $[\bar{\omega}]$. Then $\bar{\omega}$ is a merger of R.*

Partition is an operation which for all incomplete preference relations in some representation takes all the possible extensions of this relation by another set A. Merger is a reverse operation to partition.

Lemma 1. *Let U have a finite representation. Then U has a finite disjoint representation.*

Proof. I proceed by induction on m, which is the number of elements in the representation. Let $m = 1$. Then $U = [\bar{\omega}_1]$ and therefore the statement of the lemma is trivially satisfied. I just need to prove the implication that if the statement of the lemma is satisfied for some m, then it is satisfied for $m + 1$.

Assume, that for any $U = \bigcup_{i=1}^{m}[\bar{\omega}_i]$ there is a disjoint representation $U = \bigcup_{i=1}^{m'}[\bar{\omega}'_i]$. Now assume $U = \bigcup_{i=1}^{m+1}[\bar{\omega}_i]$. By assumption, there is a finite disjoint representation $R = \{\bar{\omega}'_1, \ldots, \bar{\omega}'_{m'}\}$ such that $U = \left(\bigcup_{\bar{\omega}' \in R}[\bar{\omega}']\right) \cup [\bar{\omega}_{m+1}]$.

Let $R' = \{\bar{\omega}' \in R : [\bar{\omega}'] \cap [\bar{\omega}_{m+1}] \neq \emptyset\}$ and $A = \bigcup_{\bar{\omega}' \in R} \mathrm{cp}(\bar{\omega}')$ and denote by P a partition of $\bar{\omega}_{m+1}$ by A. By definition, for each $\bar{\omega}'' \in P$ either $[\bar{\omega}''] \subset [\bar{\omega}']$ for some $\bar{\omega}' \in R'$ or $[\bar{\omega}'] \cap [\bar{\omega}''] = \emptyset$ for all $\bar{\omega}' \in R$. Let $P' = \{\bar{\omega}'' \in P : [\bar{\omega}'] \cap [\bar{\omega}''] = \emptyset\}$. Now $R \cup P'$ is a finite disjoint representation of U and the proof is complete.

□

Lemma 2. *Let $U \subset \Omega$ and fix two disjoint representations $R_0 = \{\bar{\omega}_1, \ldots, \bar{\omega}_n\}$ and $R = \{\bar{\omega}'_j : j \in \mathbb{N}_+\}$ of U such that R is subordinate to R_0. There exists a sequence $(R_l)_{l \in \mathbb{N}}$ of representations such that $\bigcap_{k=0}^{\infty} \bigcup_{l=k}^{\infty} R_l = R$ and that R_{l+1} is obtained from R_l using only partitions and mergers.*

Proof. Fix U, R_0 and R as in the statement of the lemma. I prove this lemma constructively, by providing a procedure to obtain a sequence of representations that satisfy the conditions of the lemma. For each $l \in \mathbb{N}$ I do the following steps.

Fix $\bar{\omega}_i \in R_l$, with $i = 1$ in case $l = 1$. I also fix j, starting with $j = 1$ for $l = 1$. Define $R_{\bar{\omega}_i} = \{\bar{\omega}' \in R : [\bar{\omega}'] \subset [\bar{\omega}_i]\}$ and $R_{\bar{\omega}_i}(n) = \{\bar{\omega}' \in R_{\bar{\omega}_i} : l(\bar{\omega}') = n\}$. Fix n to be the smallest number such that $R_{\bar{\omega}_i}(n) \neq \emptyset$ and denote

$$D_{\bar{\omega}_i}(n) = \bigcup_{\bar{\omega}' \in R_{\bar{\omega}_i}(n)} \mathrm{cp}(\bar{\omega}') \setminus \mathrm{cp}(\bar{\omega}_i),$$

so that $D_{\bar{\omega}_i}(n)$ is a set of points on which additional conditions in $R_{\bar{\omega}_i}(n)$ are imposed. Note, that $D_{\bar{\omega}_i}(n)$ is finite. Define $R^P_{\bar{\omega}_i}$ to be the partition of $\bar{\omega}$ with respect to $D_{\bar{\omega}_i}(n)$ and define $R^{rest}_{\bar{\omega}_i} = \{\bar{\omega}'_{|D_{\bar{\omega}_i}(n) \cup cp(\bar{\omega}_i)} : \bar{\omega}' \in R_{\bar{\omega}_i}\}$.

As both $R^P_{\bar{\omega}_i}, R^{rest}_{\bar{\omega}_i}$ are finite and disjoint and $R^P_{\bar{\omega}_i}$ is subordinate to $R^{rest}_{\bar{\omega}_i}$, I can obtain each element $\bar{\omega}' \in R^{rest}_{\bar{\omega}_i}$ by a full merger of all elements of $R^P_{\bar{\omega}_i}$ that satisfy $[\bar{\omega}''] \subset [\bar{\omega}]$. As a result, I can obtain $R^{rest}_{\bar{\omega}_i}$ from $\bar{\omega}$ using only partitions and mergers.

Finally, define $R_{l+1} = R_l \cup R^{rest}_{\bar{\omega}_i} \setminus \{\bar{\omega}_i\}$. I do not change the enumeration of elements in R_{l+1}. As such, all elements of $R_{l+1} \cap R^{rest}_{\alpha_i}$ are not enumerated (for now), meaning that there is no i' such that $\bar{\omega}_{i'} \in R_{l+1} \cap R^{rest}_{\bar{\omega}_i}$. Note that by construction $R_{\bar{\omega}_i}(n) \subset R_{l+1}$, R_{l+1} is disjoint and R is subordinate to R_{l+1}.

Now, if $\bar{\omega}_{i+1} \in R_{l+1}$, increase i, l by one and perform the same operations as I did up to this point. In the other case, increase l and j by one, set $i = 1$ and enumerate all elements of R_l.

Note, that clearly if $\bar{\omega}' \in R$ and $\bar{\omega}' \in R_l$ for any l then also $\bar{\omega}' \in R_{l+1}$. Therefore in order to finish the proof, I just need show that every element $\bar{\omega}' \in R$ is obtained as an element of R_l for some l. Let $[\bar{\omega}'] \subset [\bar{\omega}]$ for some $\bar{\omega} \in R_0$ and fix $n_{\bar{\omega}'} = |\{n \leq l(\bar{\omega}') : R_{\bar{\omega}}(n) \neq \emptyset\}|$.

I claim, that I obtain $\bar{\omega}'$ as an element of R_l for some l such that $j = n_{\bar{\omega}'}$. Indeed, if $n_{\bar{\omega}'} = 1$, I have already shown it. Consider $n_{\bar{\omega}'} > 1$. There is i such that $[\bar{\omega}'] \subset [\bar{\omega}_i]$ for $j = 1$ and there is $\bar{\omega}'' \in R^{rest}_{\bar{\omega}_i}$ such that $[\bar{\omega}'] \subset [\bar{\omega}'']$. As $R_{\bar{\omega}''} \subset R_{\bar{\omega}_i} \setminus R_{\bar{\omega}_i}(n)$ where n is the smallest number such that $R_{\bar{\omega}_i}(n) \neq \emptyset$ I get that $|\{n \leq l(\bar{\omega}') : R_{\bar{\omega}''}(n) \neq \emptyset\}| < n_{\bar{\omega}'}$, proving the claim. Since for each j I perform a finite number of partitions and mergers, the proof is complete.

□

Lemma 3. *Fix two finite disjoint representations* $R_1 = \{\bar{\omega}_1, \ldots, \bar{\omega}_{m_1}\}$ *and* $R_2 = \{\bar{\omega}'_1, \ldots, \bar{\omega}'_{m_2}\}$ *of some set* $U \subset \Omega$. *Assume that set function*

$$\mu_0 : \{U \subset \Omega : \exists_{\bar{\omega}}\ U = [\bar{\omega}]\} \to [0, 1]$$

which for any incomplete preference relation $\bar{\omega}$ *satisfies both* $\mu_0([\bar{\omega} \cup \{(x, y), (y, x)\}]) = 0$ *and* $\mu_0([\bar{\omega} \cup \{(x, y)\}]) + \mu_0([\bar{\omega} \cup \{(y, x)\}]) = \mu_0([\bar{\omega}])$ *is given. Then* $\sum_{j=1}^{m_1} \mu_0([\bar{\omega}_j]) = \sum_{j=1}^{m_2} \mu_0([\bar{\omega}'_j])$.

Proof. By the condition that $\mu_0([\bar{\omega} \cup \{(x, y)\}]) + \mu_0([\bar{\omega} \cup \{(y, x)\}]) = \mu_0([\bar{\omega}])$ the values of μ_0 are assigned in such a way that replacing any $\bar{\omega}_{j_0}$ or $\bar{\omega}'_{j_0}$ by its arbitrary partition, for example replacing $\bar{\omega}_{j_0}$ by $\bar{\omega}^1_{j_0}, \bar{\omega}^2_{j_0}$ gives $\sum_{j=1}^{m_1} \mu_0([\bar{\omega}_j]) = \mu_0([\bar{\omega}^1_{j_0}]) + \mu_0([\bar{\omega}^2_{j_0}]) + \sum_{j \neq j_0}^{m_1} \mu_0([\bar{\omega}_j])$. Therefore it suffices to show, that there exists a finite sequence of partitions from both R_1 and R_2 to some $R = \{\bar{\omega}^{l_1}_1, \ldots, \bar{\omega}^{l_1}_{k_1}\}$, meaning I can obtain the same representation as a result of the recursive partitioning of R_1 and R_2. It suffices to define

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

and fix R to be a representation obtained by partitioning of all elements of R_1 on all elements of D. Obviously, partitioning all elements of R_2 on all elements of D I also obtain R.

□

A.2 Proof

Define a family of sets

$$\mathcal{A} = \left\{ \bigcup_{j=1}^{m} \bigcap_{i=1}^{n_j} [x_{ij} R_{ij} y_{ij}] : x_{ij}, y_{ij} \in \mathcal{X}, R_{ij} \in \{\succ, \succeq, \sim\} \right\}.$$

Note that $\mathcal{A}$ is an algebra of sets. It contains an empty set as $[x \succ y] \cap [y \succ y] = \emptyset$. It obviously is closed under both binary unions and taking complements, so $\mathcal{A}$ is an algebra of sets.

I first extend μ_0 to the whole $\mathcal{A}$ as follows: define $\mu_0(\bigcup_{i=1}^{n}[x_i R_i y_i]) = 0$ if any $R_i = \sim$ and $\mu_0(\bigcup_{i=1}^{n}[x_i R_i y_i]) = \mu_0(\bigcup_{i=1}^{n}[x_i \succ y_i])$ otherwise. Moreover from proposition 1 I get that each $A \in \mathcal{A}$ has a disjoint representation, so define $\mu_0(A) = \sum_{j=1}^{m} \mu_0(\bigcap_{i=1}^{n_j}[x_{ij} \succ y_{ij}])$. Note that this is well defined due to lemma 3, which I 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.

I 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 $\bigcup_{j=1}^{\infty} A_j = A$. Moreover, denote by K the representation of A corresponding to A_j's, so that $K = \{\bar{\omega}_j : j \in \mathbb{N}_+\}$ satisfies $[\bar{\omega}_j] = A_j$. I 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 $D = \bigcup_{\bar{\omega} \in \tilde{K_0}} \text{cp}(\bar{\omega})$ and take $K_0 = \{\bar{\omega}_{|D} : \bar{\omega} \in K\}$. Clearly, K_0 is finite and K is subordinate to K_0. Therefore from lemma 2 I have that there exists 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 $l \in \mathbb{N}$

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

Now consider two sequences $m_l = \sum_{\bar{\omega} \in K^l} \mu_0([\bar{\omega}])$ 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 $\lim_{l\to\infty} m_l = \sum_{\bar{\omega}\in K} \mu_0([\bar{\omega}])$ and $\lim_{l\to\infty} m^l = \lim_{l\to\infty} \mu_0([\bigcap_{k=0}^{l} \bigcup_{k'=k}^{l} K^{k'}]) = \mu_0(\bigcup_{\bar{\omega}\in K}[\bar{\omega}]) = \mu_0(A)$. 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 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. To finish the proof it suffices to show that μ is probabilistic, but this follows trivially from the condition that $\mu_0([x \succ y]) + \mu_0([y \succ x]) = 1$. The second part of the theorem follows trivially from the first.

B Proof of theorem 2

Supporting facts and definitions

Definition 18. *I denote by* $\mathrm{Diag}(\omega), \mathrm{Diag}^+(\omega), \mathrm{Diag}_-(\omega) \subset \mathcal{X} \times \mathcal{X}$ *sets of respectively diagonal, upper diagonal and lower diagonal elements of relation* ω*, that is*

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

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

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

Definition 19. *Let* $\mu_\mathcal{K}$ *be given. I say that a measure* $\mu'_\mathcal{K}$ *is obtained from* $\mu_\mathcal{K}$ *by a disturbance* (μ', w') *if* μ' *is a probability measure defined on* $\Omega(\mathcal{K})$*, function* $w' : \mathcal{X}^2 \to [0,1]$ *satisfy* $w'(x,y) = w'(y,x)$ *and*

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

Definition 20. *Let* $\mu_\mathcal{K}$ *be given. The disturbance* (μ', w') *does not disturb the diagonal, if and only if for* $A = \mathrm{supp}(w')$ *I have all of the following*

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

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

If this is not the case, (μ', w') *disturbs the diagonal.*

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

Proof. Let $\mu'_{\mathcal{K}}$ be obtained from $\mu_{\mathcal{K}}$ without disturbing the diagonal and denote by $\omega'_{\mathcal{K}}$ (or $\succeq_{\mathcal{K}'}$) the relation given by definition 6 applied to $\mu'_{\mathcal{K}}$. Fix an arbitrary $x \in \mathcal{X}$. Following definition 20 I have $A \cap \{y \in \mathcal{X} : y \sim_{\mathcal{K}} x\} = \emptyset$. Therefore $x \sim_{\mathcal{K}} y \iff x \sim_{\mathcal{K}'} y$. Now let $y \in \mathcal{X}$ be such that $y \succ_{\mathcal{K}} x$. If $(y,x) \notin \text{supp}(w')$ then obviously $y \succ_{\mathcal{K}'} x$, so assume that $(y,x) \in \text{supp}(w')$. Now by definition of a disturbance

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

By assumption $y \succ_{\mathcal{K}} x$ I have $\mu_{\mathcal{K}}([y \succ x]) > \frac{1}{2}$. Moreover following definition 20 I have $\mu'([y \succ x]) \geq \frac{1}{2}$. Therefore $\mu'_{\mathcal{K}}([y \succ x]) > \frac{1}{2}$ and $y \succ_{\mathcal{K}'} x$. As the case with $x \succ_{\mathcal{K}} y$ is symmetric to this one, $\omega_{\mathcal{K}} = \omega'_{\mathcal{K}}$ and therefore $\mu'_{\mathcal{K}}$ also represents $\omega_{\mathcal{K}}$.

□

Proof

Due to corollary 1, I can restrict my attention only to values of $\mu_{\mathcal{K}}$ on the sets $[x \succ y]$. As $\omega_{\mathcal{K}} \in \Omega(\mathcal{K})$ 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 $\omega_{\mathcal{K}}$. As $\mathcal{X}$ is compact and $\omega_{\mathcal{K}}$ 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 $\mu_*([x \succeq y]) \geq \frac{1}{2} \iff u(x) \geq u(y)$, and therefore it represents $\omega_{\mathcal{K}}$ on Ω. Moreover, for any $z \in \mathcal{X}$ I have $\mu_*([x \succeq z]) \geq \mu_*([y \succeq z]) \iff u(x) \geq u(y)$ and therefore μ_* is coherent. However, it cannot represent $\omega_{\mathcal{K}}$ on $\Omega(\mathcal{K})$ as it is not restricted to $\Omega(\mathcal{K})$, so for $k_1 \succ k_2 \subset \bar{\omega^*}_{|\mathcal{K}}$ does not imply $\mu_*([k_1 \succ k_2]) = 1$ unless $k_1 \sim_{\mathcal{K}} x^*$ and $k_2 \sim_{\mathcal{K}} y_*$. Note however, that from definition 6 I have $k_1 \succ k_2 \subset \bar{\omega^*}_{|\mathcal{K}} \iff k_1 \succ_{\mathcal{K}} k_2$, and therefore $\mu_*([k_1 \succ k_2]) > \frac{1}{2}$.

By lemma 4, if I disturb μ_* without disturbing the diagonal, the disturbed measure also represents $\omega_{\mathcal{K}}$. I 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 $y \succeq x \in \mathcal{K}$. I can assume without loss of generality that $\mathcal{K}$ consists only of elements for which the revealed preference relation is strict, meaning that for all $k_1, k_2 \in \mathcal{K}$ I have $k_1 \succ k_2$ or $k_2 \succ k_1$ and I denote the set all known relations as $\bar{\mathcal{K}} = \{k_i \succ l_i : k_i, l_i \in \mathcal{K}, k_i \succ l_i \subset \bar{\omega^*}_{|\mathcal{K}}\}$, and $|\bar{\mathcal{K}}| = m$

For all $i \leq m$, fix some pairwise disjoint $B_i = B((k_i, l_i), r_i) \subset \text{Diag}^+(\omega_{\mathcal{K}})$ and define

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

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

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 $\mu_{\mathcal{K}}([x \succ y]) = (1 - \sum_{i=1}^{m} w_i(x,y))\mu_*([x \succ y]) + \sum_{i=1}^{m} w(x,y)\mu_i([x \succ y])$ represents $\omega_{\mathcal{K}}$. Moreover $w_i(k_i, l_i) = 1$ and $\mu_i([k_i \succ l_i]) = 1$, so μ is restricted to $\Omega(\mathcal{K})$.

C Proof of theorem 3

I start with the first point of the theorem. Let $\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}, \mu_{\mathcal{K}}([y \succeq z]) \geq \frac{1}{2}$ and denote $A = \mu_{\mathcal{K}}([z \succeq x \succeq y])$, $B = \mu_{\mathcal{K}}([x \succeq z \succeq y])$, $C = \mu_{\mathcal{K}}([x \succeq y \succeq z])$, $D = \mu_{\mathcal{K}}([z \succeq y \succeq x])$, $E = \mu_{\mathcal{K}}([y \succeq z \succeq x])$, $F = \mu_{\mathcal{K}}([y \succeq x \succeq z])$. Now $x \succeq_{\omega_{\mathcal{K}}} y \iff A + B + C \geq D + E + F$, $y \succeq_{\omega_{\mathcal{K}}} z \iff C + E + F \geq A + B + D$ and $x \succeq_{\omega_{\mathcal{K}}} z \iff B + C + F \geq A + D + E$.

From assumption that $\mu_{\mathcal{K}}([x \succeq y]) \geq \frac{1}{2}$ and $\mu_{\mathcal{K}}([y \succeq z]) \geq \frac{1}{2}$ I get that $A + B + C \geq \frac{1}{2} \geq A + B + D$ so $C \geq D$. Now assume weak coherence holds. Applying it to $x \succeq_{\omega_{\mathcal{K}}} y$ I get $B \geq E$ and applying it to $y \succeq_{\omega_{\mathcal{K}}} z$ I get $F \geq A$. Therefore I get $B + C + F \geq A + D + E$ and $\omega_{\mathcal{K}}$ is transitive. Now assume $\omega_{\mathcal{K}}$ is transitive, so $B + C + F \geq A + D + E$ holds. Due to the assumption that $x \succeq_{\omega_{\mathcal{K}}} y$ I get

$$B + C + F \geq A + D + E \iff A + B + C + 2F \geq 2A + D + E + F \implies$$

$$\implies 2F \geq 2A \iff F \geq A.$$

Similarly from the assumption that $y \succeq_{\omega_{\mathcal{K}}} z$ I get

$$B + C + F \geq A + D + E \iff 2B + C + F + E \geq A + D + 2E + B \implies$$

$$\implies 2B \geq 2E \iff B \geq E,$$

so weak coherence holds.

The second point follows trivially from the first point together with theorem 2, because if firstly, for each $\omega \in \Omega(\mathcal{K})$ there exists $\mu_{\mathcal{K}}$ representing it; secondly, each $\omega \in \Omega(\mathcal{K})$ is transitive; and finally, and for any $\mu_{\mathcal{K}}$ the expected preference relation is transitive if and only if $\mu_{\mathcal{K}}$ is weakly coherent, then for each $\omega \in \Omega(\mathcal{K})$, any conditional measure representing it must be weakly coherent.

Finally, the last point of the theorem follows trivially from the second one and corollary 1.

References

Bewley, Truman F. (2002). "Knightian decision theory. Part I". In: *Decisions in Economics and Finance* 25, pp. 79–110.

Bostic, Raphael, Richard J. Herrnstein, and Duncan R. Luce (1990). "The effect on the preference-reversal phenomenon of using choice indifferences". In: *Journal of Economic Behavior Organization* 13.2, pp. 193–212.

Butler, David J. and Graham C. Loomes (2007). "Imprecision as an account of the preference reversal phenomenon". In: *The American Economic Review* 97.1, pp. 277–297.

Carlsson, Fredrik, Morten R. Mørkbak, and Søren B. Olsen (2012). "The first time is the hardest: a test of ordering effects in choice experiments". In: *Journal of Choice Modelling* 5.2, pp. 19–37.

Cooke, Kevin (2017). "Preference discovery and experimentation". In: *Theoretical Economics* 12.3, pp. 1307–1348.

Cox, James C. and David M. Grether (1996). "The preference reversal phenomenon: response mode, markets and incentives". In: *Economic Theory* 7.3, pp. 381–405.

Czajkowski, Mikołaj, Nick Hanley, and Jacob LaRiviere (2015). "The Effects of Experience on Preferences: Theory and Empirics for Environmental Public Goods". In: *American Journal of Agricultural Economics* 97.1, pp. 333–351.

Day, Brett et al. (2012). "Ordering effects and choice set awareness in repeat-response stated preference studies". In: *Journal of Environmental Economics and Management* 63.1, pp. 73–91.

Debreu, Gerard (1964). "Continuity Properties of Paretian Utility". In: *International Economic Review* 5.3, pp. 285–293.

Engelmann, Dirk and Guillaume I. Hollard (2010). "Reconsidering the Effect of Market Experience on the "Endowment Effect"". In: *Econometrica* 78.6, pp. 2005–2019.

Gilboa, Itzhak and David Schmeidler (1995). "Case-Based Decision Theory". In: *The Quarterly Journal of Economics* 110.3, pp. 605–639.

Huang, Shiyan et al. (2014). "A new decision-making method by incomplete preferences basedon evidence distance". In: *Knowledge-Based Systems* 56, pp. 264–272.

Humphrey, Steven J., Luke Lindsay, and Chris Starmer (2017). “Consumption experience, choice experience and the endowment effect”. In: *Journal of the Economic Science Association* 3.2, pp. 109–120.

Kannai, Yakar (1970). “Continuity Properties of the Core of a Market”. In: *Econometrica* 38.6, pp. 791–815.

Kingsley, David C. and Thomas C. Brown (2010). “Preference uncertainty, preference learning, and paired comparison experiments”. In: *Land Economics* 86.3, pp. 530–544.

Kreps, David M. (1979). “A Representation Theorem for "Preference for Flexibility"”. In: *Econometrica* 47.3, pp. 565–577.

Kuilen, Gijs van de (2009). “Subjective probability weighting and the discovered preference hypothesis”. In: *Theory and Decision* 67, pp. 1–22.

Lichtenstein, Sarah and Paul Slovic (1971). “Reversals of preference between bids and choices in gambling decisions”. In: *Journal of Experimental Psychology* 89.1, pp. 46–55.

— (2006). *The construction of preference*. Ed. by Sarah Lichtenstein and Paul Slovic. Cambridge University Press.

Loomes, Graham, Shepley Orr, and Robert Sugden (2009). “Taste uncertainty and status quo effects in consumer choice”. In: *Journal of Risk and Uncertainty* 39, pp. 113–135.

Ok, Efe A., Pietro Ortoleva, and Gil Riella (2012). “Incomplete Preferences Under Uncertainty: Indecisiveness in Beliefs versus Tastes”. In: *Econometrica* 80.4, pp. 1791–1808.

Piermont, Evan, Norio Takeoka, and Roee Teper (2016). “Learning the Krepsian state: Exploration through consumption”. In: *Games and Economic Behavior* 100, pp. 69–94.

Plott, Charles R. (1996). “Rational individual behaviour in markets and social choice processes: the discovered preference hypothesis”. In: *The rational foundations of economic behaviour: Proceedings of the IEA Conference held in Turin, Italy.* Ed. by Kenneth J. Arrow. Vol. 114. New York: St. Martin’s Press; London: Macmillan Press in association with the International Economic Association, pp. 225–250.

Plott, Charles R. and Kathryn Zeiler (2005). “The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations”. In: *American Economic Review* 95.3, pp. 530–545.

Stalmeier, Peep F.M., Peter P. Wakker, and Thom G.G. Bezembinder (1997). “Preference reversals: Violations of unidimensional procedure invariance”. In: *Journal of Experimental Psychology: Human Perception and Performance* 23.4, pp. 1196–1205.

Sugden, Robert (2003). “Reference-dependent subjective expected utility”. In: *Journal of Economic Theory* 111.2, pp. 172–191.

Tversky, Amos, Shmuel Sattath, and Paul Slovic (1988). “Contingent weighting in judgment and choice”. In: *Psychological Review* 95.3, pp. 371–384.