QUEEN MARY UNIVERSITY OF LONDON

Essays on Stochastic Choice and Welfare

by
Daniele Caliari

Thesis submitted in partial fulfillment of the requirements of the
Degree of Doctor of Philosophy (Ph.D.) in Economics

in the
School of Economics and Finance
Queen Mary University of London

May 4, 2021

Declaration of Authorship

I, Daniele Caliari, confirm that the research included within this thesis is my own work or, where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Abstract

Motivated by the literature on preference elicitation and welfare analysis, Chapter I studies the properties of aggregators of choice datasets into preferences. Novel normative principles and their theoretical implications are provided. I analyse numerous approaches proposed by the literature in view of the introduced principles. I also propose and characterize two counting procedures that are foundational for the analysis.

Motivated by the theoretical framework of the first chapter, in Chapter II, I propose a novel experimental design to test two normative principles: (1) Informational Responsiveness guarantees that no choice data is ignored; (2) Revealed Preference constrains the preference elicitation process to a particular reorganization of data. These principles are summarized by a method denoted as Counting Reveal Preference procedure. I show that approaches founded on this procedure provide more reliable results in terms of preference relation.

Motivated by the literature on stochastic choice, Chapter III studies the relation between imperfect discrimination and the transitivity of preferences. I show that the degree of transitivity depends on the degree of discrimination between pairs of alternatives. I characterize the notions of Weak, Moderate and Strong stochastic transitivity. The results allow us to organize a wide range of stochastic models in accordance with Fechnerian models and imperfect discrimination.

Acknowledgements

This essay is the partial outcome of five years of mistakes. The journey to the current version of this thesis had been characterized by a surprisingly high number of badly written notes and drafts. Note that this does not imply that the current version is in any way well-written or clear. I am inclined to say that the contrary is more likely to be true. This may inform the reader regarding the quality of my writing and logical thinking at the beginning of the journey.

The previous paragraph is the premise for what has to come. Among the many, for the improvement in the way I can think and describe my ideas, I thank my first supervisor, Professor Marco Mariotti. The patience he had, when listening to the numerous wrong ideas that I was able to come up with, has been frankly remarkable. I tried to exchange his suggestions with some football tips or mountain photos but I probably have to admit that we are far from even.

I thank my second supervisor, Dr. Christopher Tyson, for the numerous discussions regarding some of the technical results in this thesis. I also had the fortune to assist Dr. Tyson in teaching his undergraduate Game Theory course. He is surely the best teacher I have ever seen and an example of clarity, organization and dedication to the job of teaching, whose importance is too often forgotten. I also thank my supervisors at the University of Verona, Professor Maria Vittoria Levati and Dr. Ivan Soraperra, for convincing and guiding me to apply for a Ph.D. Looking backward, it has been a fun experience.

I say sorry to my colleagues and friends for the loss in productivity due to my desire in arguing topics that have no relation whatsoever with my or their topics of research. In my defence, I followed Christopher Hitchens's suggestion: "Time spent arguing is, oddly enough, almost never wasted."

Finally, if I was home with my parents and sister, this would be the time when we start arguing about why I should thank my family. The argument: "everybody does it" seems weak. Furthermore, not going into details would appear to the reader as if I follow exactly that argument. Hence, I will try to be slightly more specific in defining the reasons. One of them, my mother would say I should not mention. Therefore, I will not write "money", but financial support. However, the main reason lies in a policy that is largely and luckily applied in my family, that is: "Fa quel che te vol che te vive pu en pez" or the english version "Do whatever you want that you are going to live longer". So, I thank my family for the absence of restrictions in the decisions I took along the way.

Contents

List of Figures

List of Tables

Introduction

This dissertation comprises three chapters organized as follows: Chapter I and II study Behavioural Welfare Analysis from a theoretical and empirical point of view. Throughout these chapters, numerous microeconomic topics are spanned: deterministic and stochastic choice theory, aggregation theory, welfare theory, choice-elicitation experimental design. Chapter III studies Choice Behaviour under Imperfect Discrimination. It spans two topics: stochastic choice theory and order theory.

Chapter I aims to reconcile welfare analysis with non-standard choice behaviour. It is motivated by the following example: suppose a researcher wants to elicit preferences from a group of individuals' choices. They may behave inconsistently with utility maximization. Therefore, the researcher faces the challenge of identifying a preference relation for each of these individuals. As a further complication behaviour is heterogeneous. Namely, each individual violates utility maximization differently. As a consequence, the researcher has to potentially identify numerous different (correct) models, one for each individual. The literature has acknowledged the complexity of this problem and proposed different solutions. An influential approach is due to Bernheim & Rangel (2009) and suggest to ignore inconsistent data and apply standard theory on the remaining dataset. This approach highly simplifies the problem, however at a considerable price: the elicited preferences may be very coarse. I propose a theoretical solution that is a refinement of Bernheim & Rangel (2009) and that is based on two requirements: an axiom called Informational Responsiveness and a notion of frequency based on Revealed Preference. I, then, consider numerous other requirements and completely characterize two notions of frequency with datasets that may have both multiple observations and missing data.

Chapter II brings the theoretical results of the first chapter to the data. I design a novel choice-elicitation experiment. Two questions, that constitute the premise and test of my theoretical study, are addressed: do subjects behave consistently with utility maximization and do they exhibit heterogeneous modes of behaviour? If individual choices present inconsistencies, do my requirements improve the capacity of the researcher to elicit preference relations? To both questions the answer is affirmative. I find that subjects show high inconsistent behaviour both in Time and Risk preferences. Furthermore, many different behavioural models are adopted. I find that even within the same subject, different modes of behaviour are displayed moving from Time to Risk environments. These findings pose a serious challenge to the researcher who wants to elicit preferences. However, I show that when the requirements studied in Chapter I are satisfied, the preference elicitation capacity increases significantly. I test the results both indirectly analysing approaches proposed by the literature and directly creating a novel test for Informational Responsiveness.

Chapter III studies the topic of Imperfect Discrimination in relation to transitivity of preferences and stochastic choice. I study how imperfect discrimination, namely the inability to distinguish the utility of stimuli, have effects into transitivity of choices and binary relations. I build on threshold (Aleskerov et al., 2007) and perturbed models (Fudenberg et al., 2015) to precisely characterize the connection between imperfect discrimination and transitivity. As the main result, I characterize all notions of stochastic transitivity (Weak, Moderate and Strong) using only a parameter η that measures the degree to which an individual departs from standard utility maximization. I show that a sufficient condition is related to the triangle inequality. Finally, I provide a connection between deterministic and stochastic transitivity completing a study initiated by Fishburn (1973).

Chapter 1

Behavioural Welfare Analysis and Revealed Preference: Theory

1.1 Introduction

In recent years behavioural economics has developed a large number of models in response to evidence of violations of the standard model of decision-making. This growing literature raises the problem of selecting behavioural models to analyse datasets and contribute to policy evaluation. More concretely, imagine a researcher that possesses a choice dataset from multiple individuals that choose following different behavioural models. For instance, some are perfectly rational, others may face cost of thinking (Ortoleva, 2013), (Fudenberg et al., 2015), (Frick, 2016), form consideration sets (Manzini & Mariotti, 2014), (Brady & Rehbeck, 2016), use attention filters (Masatlioglu et al., 2012), (Lleras et al., 2017), (Cattaneo et al., 2018), perception orders (Echenique et al., 2018), checklists (Mandler et al., 2012) or sequential rationales (Manzini & Mariotti, 2007). The researcher is interested in eliciting individual preferences. Therefore, in each case, she has to identify the correct model and elicit the preference relation accordingly.

The literature has acknowledged the complexity of the researcher's task. Some choices may not be in line with any model at disposal while others may be in line with more than one. Some individuals may also be endowed with more than one model and exchange them according to different situations. An influential approach, that can be applied irrespectively of the model, and therefore overcome these difficulties, is due to Bernheim & Rangel (2009). The authors develop a Pareto relation that cautiously considers **x** preferred to **y** if and only if **y** is never chosen when **x** is available. With this in mind, the researcher's task is simplified. She does not need to identify the different models. She has only to observe their choices and apply the above definition. However, the simplification comes at a considerable price: the elicited preference may be very coarse. This problem has been highlighted by Rubinstein & Salant (2012) as follows: "The resulting Pareto relation is typically a coarse binary relation that becomes even more so as the behavioural dataset grows".

Bernheim & Rangel (2009) were aware of the coarseness problem. Therefore, they propose some reasonable refinements, each accompanied by some drawbacks. The first proposal is related to the availability of further information. If the researcher possesses data on imperfect information processing or clear mistakes she can use them to enrich the choices. However, this is not always possible and involves subjective decisions. The second proposal is to carefully consider only part of the dataset for reasons of (i) simplicity or (ii) frequency. In the first case, the researcher may think that choices from simpler problems (e.g. binary sets) are more reliable in eliciting preferences. In the second case, she may think that an alternative that is chosen more often is more probable to be preferred. As the authors notice, both approaches leave the researcher with the issue of choosing which problems are simpler or which definition of frequency to choose.

In this chapter, we propose a solution to the researcher's problem that is a refinement of Bernheim & Rangel (2009) approach. We acknowledge Rubinstein & Salant (2012) critique and propose a normative principle that guarantees that "more data lead to finer results".

Our normative principle states that a researcher that aims to rank two alternatives **x** and **y** has to use all the relevant evidence about **x** and **y**. To formalize, we say that when she considers **x** indifferent to **y**, more choices of **x** with **y** available should turn the judgement in favour of **x**.[1] We call this condition Informational Responsive-

[1]Note that this requirement can be weakened without losing its interpretation. This feature is in line with Krantz et al. (1971): "One demand is for the axioms to have a direct and easily understood meaning in terms of empirical operations, so simple that either they are evidently empirically true on intuitive grounds or it is evident how systematically to test them." In our case, the reader may think

ness.[2] A violation would imply that she is not using all relevant data. Notice also that we are adopting a particular notion of frequency. That is, we consider relevant only the data that shows one element chosen when the other is also available, as in standard revealed preference.

But, how can two such simple conditions provide a reliable solution to the complex researcher's problem? The answer lies in the monotonic connection between preferences and choices that most models share. Behavioural models are made of different components (e.g. mistakes, cost of thinking, alternatives attributes, lists of rationales, perception orders, and many others). However, all of them are monotonic in the underlying preference relation. Even models that seem to depart from this principle, such as attention models, satisfy the property that $\mathbf{x}$ is chosen more times than $\mathbf{y}$ if $\mathbf{x}$ is preferred to $\mathbf{y}$ and attention and preferences are not too conflictual.[3] The monotonic component of behavioural models has been highlighted by Apesteguia & Ballester (2015). The authors show that their approach, that satisfy our conditions, constitutes a reliable solution to the researcher's problem when individuals adopt one of a broad family of models summarized by a property that they call P-Monotonicity.[4]

We generally define the researcher's problem as behavioural welfare analysis (henceforth BWA). More formally, BWA maps individual choices into welfare orderings (binary relations) and aims to deal with non-standard patterns of choice. To relate with the classic literature, standard welfare analysis is the subset of BWA based on the assumption that choices satisfy the Weak Axiom of Revealed Preference [WARP][5] and therefore that they are revealed to be maximized by a transitive and complete preference relation (Sen, 1971). In this case, the welfare ordering is identified with the maximized preference relation.[6] The literature has exten-

of requiring more than choice to break the indifference.

[2]To the best of our knowledge this property has been firstly introduced in voting theory by Goodin & List (2006) under the denomination of "One Vote Responsiveness".

[3]For example, in Manzini & Mariotti (2014), if $\mathbf{x}$ is preferred to $\mathbf{y}$ then in order for $\mathbf{y}$ to be chosen more frequently than $\mathbf{x}$, the attention parameter (salience) of $\mathbf{y}$ has to be significantly higher than the one of $\mathbf{x}$.

[4]These results are summarized in Proposition 1 and Theorem 1 of Apesteguia & Ballester (2015).

[5]A non-formal definition of WARP is as follows: If an alternative $\mathbf{x}$ is chosen when $\mathbf{y}$ is available then $\mathbf{y}$ is not chosen when $\mathbf{x}$ is available.

[6]An analogous argument can be applied to a stochastic choice that satisfies Independence from Irrelevant Alternatives (Marschak & Block, 1960), (Luce, 1959). Note that, a stochastic choice function is a refinement of a standard choice function where for each set we can observe the frequency of choice

sively documented that individuals violate not only WARP (Echenique et al., 2011) but also Independence from Irrelevant Alternatives (Tversky & Russo, 1969) and weaker assumptions such as weak stochastic transitivity (Tversky, 1969) and regularity (Huber et al., 1982), (Iyengar & Kamenica, 2010). Henceforth, we will refer to each map constituting BWA as welfare methods. Informational Responsiveness, Revealed Preference, and the other principles (or guidelines) that will be introduced along the chapter come as restrictions on the welfare methods.

Focusing on Informational Responsiveness, we argue that its desirability as a necessary condition is related to its weakness, non-triviality,[7] and relevance. We show that some very different welfare methods satisfy this axiom (*weakness*). However others do not (*non-triviality*). In such cases, we show that the violation can potentially lead to paradoxical results (*relevance*). Particularly, for a broad family of stochastic models of choice, Informational Responsiveness is necessary to infer the underlying deterministic utility. As a consequence, some methods that do not satisfy our requirement, such as (Bernheim & Rangel, 2009) and transitive core (Nishimura, 2017) fail to infer the utility function of such models no matter how large the dataset is.

Although, Informational Responsiveness keeps the spotlight; we propose other normatively appealing axioms. We introduce two continuity requirements: *stability* and *robustness* (note that no necessity is claimed here). In words, if the researcher judges $\mathbf{x}$ to be better than $\mathbf{y}$ then a single piece of evidence in favour of $\mathbf{y}$ cannot reverse the judgement (stability) and, if she judges $\mathbf{x}$ to be "much" better than $\mathbf{y}$, then again a single piece of evidence cannot make $\mathbf{y}$ either equally good or better than $\mathbf{x}$ (robustness).

Continuity requirements are normatively of particular interest because they highlight an intrinsic contrast between standard revealed preference analysis and the widely accepted idea of choice overload (Iyengar & Kamenica, 2010), (Fudenberg et al., 2015), (Frick, 2016), cost of thinking (Ortoleva, 2013), rational inattention (Matejka & McKay, 2015) and considerations sets (Manzini & Mariotti, 2014). We

for each alternative.

[7]We use the term "trivial" in relation to its logic definition. In particular, we intend as "trivial", an axiom that is satisfied by every method proposed by the literature and therefore always true in the discipline.

show how methods based on revealed preference may assign an excess weight on data from large sets (Section 1.6.2). If these observations are noisier than those from binary sets then it is unclear how they should be considered in terms of welfare revelation.[8] In Chapters II this question will be addressed empirically and answered negatively. Standard revealed preference will be shown to be crucial to elicit preferences and therefore the continuity requirements, as here defined, to be too demanding.

This chapter locates in the axiomatic approach recently proposed by Nishimura (2017) and Horan & Sprumont (2016). However, it differs from both of them. From the former because our primitive are choice observations and not preference relations. From the latter because, as a novelty, our axioms deal with the problems of information and continuity. Furthermore, our approach allows for greater flexibility in the structure of the dataset in terms of both missing data and multiple observations.

Overall we provide three main theoretical results:

- We show that *Informational Responsiveness* is the key axiom that gives rise to a class of methods that can infer the deterministic utility underlying i.i.d. Random Utility Models (Proposition 1 and 2). The results can be easily generalized to a broader family of stochastic models.

- We provide a characterization of the Counting Choice Method (the best alternative is the one chosen most times; the second best is the one chosen second most times and so on). Although this method seems naive, it constitutes an important benchmark for our theoretical analysis. Furthermore, even though counting procedures have been extensively studied a complete characterization in the context of choice datasets is a novelty (Theorem 2).

- We provide a characterization of the Counting Revealed Preference Method (if $\mathbf{x}$ is chosen when $\mathbf{y}$ is available more times than $\mathbf{y}$ when $\mathbf{x}$ available then $\mathbf{x}$ is better than $\mathbf{y}$) - (Theorem 3). We show that this method has at least two appealing properties: (1) in certain cases it is equivalent to the Minimum Swaps

[8] The definition of revealed preference plays a key role in this argument. An approach based only on binary sets has been proposed by Arrow (1959) and Sen (1971).

Method (Apesteguia & Ballester, 2015) - (Theorem 1) - which is related to a well-known computational problem (Dean & Martin, 2015); (2) it can be used as the foundation for other methods such as Eigenvector Centrality Method or Transitive Core Method (Nishimura, 2017).

Our second and third results are connected with two axiomatizations of counting methods in the class of tournament (Rubinstein, 1980) and directed graph (van den Brink & Gilles, 2003). However, our characterization differs from the above. Firstly, because it is the first application in the context of choice menus where we show that a simple monotonicity requirement as in Rubinstein (1980) is not sufficient.[9] Secondly, because our definition of welfare methods does not require the resulting binary relation to be transitive, but only to be complete and reflexive. Dropping transitivity is necessary to compare the two characterizations given that the revealed preference relation can easily be cyclic.[10] Importantly we show that the difference between our results and Rubinstein (1980) is not due to transitivity since the non-sufficiency of monotonicity is proven without any reference to transitivity.

The reader, in view of the proposal of Bernheim & Rangel (2009), may see the dropping of the requirement of acyclicity of the welfare relation as problematic. As we show, there is a trade-off between the requirement of acyclicity of the revealed preference relation and the coarseness of the relation itself. In Section 1.6, in particular, we show how the proposal of Bernheim & Rangel (2009) imposes coarseness almost everywhere through acyclicity. Differently from the ex-ante imposition of acyclicity, we provide a series of methods that allows to break the cycles and re-establish the acyclicity of the welfare relation ex-post. In Chapter II, we will show that this approach allows for a better elicitation of the welfare relation.

1.1.1 Structure of the Chapter

Section 1.2 introduces the general framework. In Section 1.3, we present Informational Responsiveness as the main conceptual axiom and argue about its necessity due to its weakness and relevance. In Section 1.4 there is a description of the

[9] A tournament is an asymmetric binary relation that can be also described using solely binary sets.

[10] A classical example: A is chosen from $\{A, B\}$, B is chosen from $\{B, C\}$ and C is chosen from $\{A, C\}$.

methods analysed in Chapter I and II together with the above mentioned equivalence result between counting revealed preference and the minimum swaps method (Apesteguia & Ballester, 2015). In Section 1.5 we show the non-triviality of Informational Responsiveness. Section 1.6 deals with the problem of sensitivity of methods (continuity requirements). Finally, in Section 1.7 we propose the characterizations of the two counting procedures: the counting choice method and the counting revealed preference method. Auxiliary results and proofs, where not contained in the text, are in Appendix A.

1.2 Framework

1.2.1 Dataset

Let X be a finite set of alternatives. Let $\mathcal{X}$ be the set of all non-empty subsets of X. Assign to any set $S \in \mathcal{X}$ a non-negative integer n that denotes the number of times the set S is observed in the data. The set of observed $S-$sets is denoted as S_n. A domain D is the collection of all the $S-$sets. Formally, $D = \bigcup_{S \in \mathcal{X}} S_n$. We denote as $\mathcal{D}$ the set of all possible domains. The domain $D_\emptyset$ denotes an empty dataset.

A choice function is then defined as $C : D \to X$ s.t. $C(S) \in S$ for all non-empty $S \in D$. Denote $\mathscr{C}(D)$ as the set of choice functions over a given D. For simplicity we denote C_D any $C \in \mathscr{C}(D)$. Let $\mathscr{C}$ be the set of all choice functions over all possible domains.

A dataset is a tuple (D, C) where D is a domain and C is a choice function defined over the domain. We define two counting measures:

1. C_x is the number of times an element $x \in X$ is chosen from any set. Formally, $C_x = |\{S \in D : x = C(S)\}|$.

2. C_{xy} is the number of times an element x is chosen when y is available. Formally, $C_{xy} = |\{S \in D : x = C(S), y \in S\}|$.

1.2.2 Welfare method

Let $\mathcal{B}(X)$ be the set of strict total orders defined over X (denoted as P^*), let $\mathcal{R}(X)$ be the set of complete, reflexive but not transitive binary relations and $\mathcal{T}(X)$ the set of transitive, reflexive but not complete binary relations (as usual we refer to P as the asymmetric part and to I as the symmetric part).

A welfare method, or simply a *method*, maps choice functions into binary relations. Hence, a method can be defined differently with respect to the class of binary relations in the codomain. Throughout the chapter, we use three definitions:

1. A method f is a correspondence $f : \mathscr{C} \rightrightarrows \mathcal{B}(X)$ and $\mathcal{F}$ is the family of all these methods. Given a generic method $f \in \mathcal{F}$, a choice function C and two elements $x,y \in X$ then if $(x,y) \in P^* \in f(C)$ we write $xP^*_f(C)y$.

2. A method g is a function $g : \mathscr{C} \to \mathcal{R}(X)$, let's denote the family of these methods as $\mathcal{G}$. Given a generic method g, a choice function $C \in \mathscr{C}$ and two elements $x,y \in X$ then if $(x,y) \in R = g(C)$, we write $xR_g(C)y$ (we write $xP_g(C)y$ for the asymmetric part and $xI_g(C)y$ for the symmetric part).

3. A method t is a function $t : \mathscr{C} \to \mathcal{T}(X)$. We denote the family of these methods as $\mathcal{T}$.

The welfare methods introduced by the literature are covered by these three families $\mathcal{F},\mathcal{G},\mathcal{T}$. In order to compare them, we define one methodology that will be extensively used throughout Chapters I and II. If for all $C \in \mathscr{C}$ and $g(C)$, P_g is acyclic, then we can connect the families of methods $\mathcal{F},\mathcal{G}$ in the following way: suppose $R = g(C)$ such that for $x,y \in X$, $(x,y),(y,x) \in R$ then we can rewrite R as two distinct $P^*_1, P^*_2 \in f(C)$ with $(x,y) \in P^*_1$, $(y,x) \in P^*_2$. In other words, indifferences in R are broken using two strict total orders P^*_1, P^*_2. If P_g is cyclical we adopt the convention of substituting the cycles with indifferences. Conversely, suppose $P^*_1, P^*_2 \in f(C)$ such that $(x,y) \in P^*_1$ and $(y,x) \in P^*_2$ then we can rewrite two P^*_1, P^*_2 as a single $R \in g(C)$ such that $(x,y),(y,x) \in R^*$.

The axioms and definitions in the following sections are defined over a complete and reflexive binary relation $R \in g(C_D)$. By abuse of notation, we will denote $R^D_g(C)$ as R^{C_D}; $R^{D\cup\{S\}}_g(C)$ as $R^{C_{D\cup S}}$ and $R^{D\setminus\{S\}}_g(C)$ as $R^{C_{D\setminus S}}$ for a generic set $S \in D$.

It is crucial to remember the reader that the axiom of Completeness plays a decisive role in all the results. However, since it is embedded into the definition of g methods, it won't be explicitly recalled in the statements.

Axiom 1 (*Completeness*).

$$\text{For any } x, y \in X, \text{ either } xR^{C_D}y \text{ or } yR^{C_D}x.$$

1.3 Informational Responsiveness

This property guarantees that a method that ranks two alternatives $x, y \in X$ uses all the relevant choice observations regarding x and y. In the case of choice-based welfare analysis it is natural, and hardly questionable, to consider relevant for x, y at least those observations where x is chosen and y is available or vice versa. This axiom tests if an observation is "key" to solve indifferences. In other words, if a method ranks x indifferent to y, then an additional observation carrying x chosen and y available should make the method rank x better than y. If this does not happen, we infer that the method is not using that observation[11].

Shortly, given a method $g \in \mathcal{G}$ and a choice function C, for all $x, y \in X$ and $D \in \mathcal{D}$:

Axiom 2 (*Informational Responsiveness* **[IR]**).

$$\text{If } xI^{C_D}y \text{ and } x = C(S),\, y \in S \text{ then } xP^{C_{D\cup S}}y. \text{ If also } S \in D \text{ then } yP^{C_{D\setminus S}}x.$$

The definition of Informational Responsiveness must care specifically about adding and removing a piece of information. The necessity is due to the extreme weakness of the antecedent ($xI^{C_D}y$) that doesn't allow to guarantee an equivalence. In Appendix A.1.2, in the proof of Claim 1, we provide a counterexample that shows the independence of adding and removing data.[12]

1.3.1 Weakness and relevance of Informational Responsiveness

We argue that Informational Responsiveness should be a necessary condition for welfare methods. We show that it has two characteristics that are desirable for a

[11]This idea is not new in the literature. In a totally different context, a similar approach has been used by Dekel et al. (2001) in order to give a definition of "relevant" state of the world.

[12]It is interesting to see that such a requirement is redundant for an axiom called *(Strong) Positive Responsiveness*, which is similar to the one proposed by Rubinstein (1980):

Axiom. *If* $xR^{C_D}y$ *and* $x = C(S)$ *then* $xP^{C_{D\cup S}}y$.

It is immediate to see the equivalence between this definition and the following: if $xR^{C_D}y$ and $y = C(S)$ then $xP^{C_{D\setminus S}}y$.

necessary axiom: *weakness* and *relevance*. Namely, it is crucial to avoid paradoxical results (relevance), but the restriction it imposes on the family of methods is not strong enough to identify even indisputable welfare relation (weakness, i.e. x always chosen and y never chosen). This task requires the introduction of two axioms: *Neutrality*[13] and *Choice non-negativeness*. The first requires that welfare analysis does not depend on the label of the alternatives; the second requires that choices are not negative evidence of the goodness of the alternatives.

Let $\Pi(X)$ be the set of all the permutations $\pi : X \to X$. Then, for all $\pi \in \Pi(X)$, define $\pi(C_D) \in \mathscr{C}(D)$ as $\pi(C(S)) := \pi(C(\pi^{-1}(S)))$ for all $S \in D$.

Axiom 3 (*Neutrality* **[NEU]).**

$$xR^{C_D}y \textit{ if and only if } \pi(x)R^{\pi(C_D)}\pi(y) \textit{ for all } \pi \in \Pi(X).$$

Axiom 4 (*Choice non-negativeness* **[CNN]).**

$$\textit{If } xI^{C_D}y \textit{ and } x = C(S) \textit{ then } xR^{C_{D\cup S}}y. \textit{ If also } S \in D \textit{ then } yR^{C_{D\setminus S}}x.$$

$$\textit{If } xP^{C_D}y \textit{ and } x = C(S) \textit{ then } xP^{C_{D\cup S}}y.$$

$$\textit{If } xP^{C_D}y, S \in S \textit{ and } y = C(S) \textit{ then } xP^{C_{D\setminus S}}y.$$

One could note that Choice non-negativeness and Informational Responsiveness together provide a sort of monotonicity (*Positive Responsiveness* - Rubinstein (1980)) over the sets S where $x = C(S)$ and $y \in S$. However, we split them for two reasons: (1) Choice non-negativeness, unlike Informational Responsiveness, is satisfied by all methods proposed by the literature; (2) Choice non-negativeness doesn't provide any insights about the informational capacity of welfare methods.

The role of the monotonicity provided by Informational Responsiveness and Choice non-negativeness has an important impact on preference elicitation. We analyse monotonicity more in details in Section 1.8. Here, however, it is interesting to note that several behavioural models proposed by literature can be captured by our monotonicity assumption as noted by Apesteguia & Ballester (2015). In the results that follow, we focus on one specific model, however, as we will discuss

[13]As highlighted by Apesteguia & Ballester (2015), Neutrality can be considered trivial in an abstract setting where there is no additional information on the alternatives. In the case where such information is available (e.g. monetary values, attributes, etc...) one could prefer to treat alternatives differently.

in Section 1.4 introducing the counting revealed preference procedure, the results hold at a higher generality level. Even models that generally do not satisfy our monotonicity assumption, such as attention models, may be captured if the tension between preferences and attention parameters is not too high (Manzini & Mariotti, 2014).

To show that Informational Responsiveness avoids paradoxical results we consider a case in which the resulting preference order is indisputable and show that it can be inferred only by methods that satisfy such property.

Thus, we introduce Random Utility Models with independent and identically distributed error components as follows. Suppose an individual evaluates the alternatives according to a utility function $u : X \to R_{++}$. However, at the act of choice this utility is perturbed by an additive error component such that the choice depends on the random utility $U(x) = u(x) + \epsilon(x)$ where $\epsilon(x)$ is continuously distributed. The probability that x is chosen from a set $S \in D$ is $Pr[x = \mathsf{argmax}_{x \in S} U(x)]$.

Furthermore, suppose that the collection of observations is restricted to multiple observations over a single set S such that $D = S_n$. We show that given this particular restriction on the domain, our three axioms can correctly identify the underlying deterministic utility u and consequently the correct welfare relation.

Proposition 1. *Given an i.i.d. RUM, a resulting collection of observations on a domain of the type $D = S_n$ and a method g that satisfies Informational Responsiveness, Neutrality and Choice non-negativeness then $xR_g y$ if and only if $u(x) \geq u(y)$.*

Proof. Since the collection of observations is produced by an i.i.d. RUM, the following clearly holds: $C_x \geq C_y$ if and only if $u(x) \geq u(y)$ when the number of observations is large. The *only if* part is trivial. Hence, we will prove only the *if* part.

Given two generic elements $x, y \in X$ we can divide the collection of observations over the domain $D = S_n$ in three disjoint sets with the following cardinality: C_x, C_y and C_z where this latter is defined as: $C_z = \sum_{z \neq x,y} |\{S \in D : z = C(S)\}|$. First, focus on this latter set; by Neutrality we have $xI^{C_D}y$. Suppose $xP^{C_D}y$; then take $\pi(x) = y$, $\pi(y) = x$ and $\pi(z) = z$ for all z, then we have $yP^{C_D}x$ but the collection of observations hasn't changed contradicting the definition of method as single-valued function. Then, take the sets of observations where x, y are chosen. The proof is by induction on $C_x + C_y$. The inductive base is proved for $C_x + C_y = 2$. First suppose $C_x + C_y = 0$ then $xI^{C_D}y$ by Neutrality. If $C_x + C_y = 1$ and x is chosen, then by Informational Responsiveness and Neutrality $xP^{C_D}y$. If $C_x + C_y = 2$ and $C_x > C_y$ then $xP^{C_D}y$ by Choice non-negativeness; if $C_x = C_y$ then $xI^{C_D}y$ by Neutrality. Suppose this statement holds for $C_x + C_y = n$ and add an additional observation such that $D = S_n \cup T$ and $x = C(T)$ (we don't need to analyse the case if $y = C(T)$, since the result would hold by definition of method as a function). Then if $C_x - C_y = 1$, $xP^{C_D}y$ by Informational Responsiveness and the inductive hypothesis; if $C_x - C_y > 1$, then $xP^{C_D}y$ by Choice non-negativeness and the inductive hypothesis. Finally, if $C_x = C_y$ then $xI^{C_D}y$ by Neutrality. □

The reader may note that the weakness of Informational Responsiveness comes not only from the use of Choice non-negativeness and Neutrality but also from the strong restriction imposed on the domain. This restriction is indeed extremely severe. Hence, a similar result is proven for a larger set of domains at the cost of requiring the resulting binary relation to be transitive. Nonetheless, a weaker restriction has to be maintained. Particularly, a domain D is *homogeneous* if it assigns to all $S \in \mathcal{X}$ the same natural number n. Equivalently, D is homogeneous if any non-empty subset is observed the same, large enough, number of times.

Proposition 2. *Given an i.i.d. RUM, a resulting collection of observations over a homogeneous domain and a method g that satisfies Informational Responsiveness, Neutrality, Choice non-negativeness and Transitivity then xR_gy if and only if $u(x) \geq u(y)$.*

Proof. See Appendix A.1.1 □

The following example shows the independence of Transitivity from the other axioms.

Example 1. *Let $D_{xy} = |\{S \in D : z = C(S), x \in S, y \notin S\}|$. Define the following method:*

$$F_{xy} \geq F_{yx} \Leftrightarrow xR^C y$$

where $F_{xy} = \delta \cdot C_{xy} + D_{xy}$ with $\delta \in \Re^{++}$.

This method satisfies Informational Responsiveness, Neutrality, Choice non-negativeness but not Transitivity. If δ is small then we have $yP^{C_D}x$ with $u(x) > u(y)$ on a homogeneous domain. For instance, suppose we observe 60 choices from $\{x,y\}$, $\{y,z\}$, $\{x,z\}$, $\{x,y,z\}$. Furthermore, suppose $u(x) = 3$, $u(y) = 2$, $u(z) = 1$ and the decision maker follows a standard Luce Model. Then, the following dataset is observed (note that (n,m) from $\{x,y\}$ indicates that x is chosen n times and y is chosen m times):

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{y,z\} & \{x,z\} \\ C(S) & (30,20,10) & (36,24) & (40,20) & (45,15) \end{array}\right]$$

$C_{xy} = 66$, $C_{yx} = 44$, $D_{xy} = 15$, $D_{yx} = 20$. *Setting δ small gives $F_{xy} < F_{yx}$. In this example, for instance, setting $\delta = 0.3$: $zP_g^{C_D}xP_g^{C_D}yI_g^{C_D}z$ violating transitivity.*

The reader may note that this example relies on a very unusual method. In fact, the axiom of Transitivity has a limited role in proving the result and can be substituted for instance by the axiom of Independence:

Axiom 5 (*Independence*)**.**

$$\textit{For all } S \in D \textit{ if } z = C(S) \textit{ then } xR^{C_D}y \Leftrightarrow xR^{C_{D\setminus S}}y.$$

The tight restriction on the domain together with the assumption of choice observations based on i.i.d. RUMs guarantee that most methods are in fact transitive on this subspace of $\mathscr{C}$. An example is the method based on C_{xy}, which is clearly not transitive over $\mathscr{C}$. Importantly, both Bernheim & Rangel (2009) [Theorem 1] and Apesteguia & Ballester (2015) [Theorem 1] implicitly rely on the restriction of homogeneous domains.

Corollary 1. *Given an i.i.d. RUM, a resulting collection of observations over a homogeneous domain and a method g that satisfies Informational Responsiveness, Neutrality, Choice non-negativeness and Independence then xR_gy if and only if $u(x) \geq u(y)$.*

1.4 An overview of *methods*

This section contains concise descriptions of the methods that will be analysed in the following sections and in Chapter II. A reader interested in specific results can skip the section and eventually refer to it at a later time.

The methods are denoted as follows: **CC** $\in \mathcal{G}$ is the *counting choice method*, **CRP** $\in \mathcal{G}$ is the *counting revealed preference method*, **SEQ** $\in \mathcal{F}$ is the *sequential method* (Horan & Sprumont, 2016), **BR** $\in \mathcal{G}$ is the *Bernheim, Rangel method* (Bernheim & Rangel, 2009), **MS** $\in \mathcal{F}$ is the *minimum swaps method* (Apesteguia & Ballester, 2015), **EIG** $\in \mathcal{G}$ is the *eigenvector centrality method* and **TC** $\in \mathcal{T}$ is a variation of the *transitive core method* (Nishimura, 2017).

Counting choice

For all domain $D \in \mathcal{D}$, the counting choice method **CC** $\in \mathcal{G}$ is simply defined as follows:

$$xR^{C_D}_{\mathbf{CC}}y \text{ if and only if } C_x \geq C_y$$

Counting revealed preference

So far we haven't assumed neither acyclicity nor transitivity defining the class of methods $\mathcal{G}$. The reason is that this allows us to apply the counting procedure to the standard revealed preference relation as a method of type g.[14] Its inclusion in $\mathcal{G}$ is driven by the following arguments: (1) **CRP** is the foundation for other important methods such as **MS**, **EIG** and **TC**; (2) the acyclicity of $P^{C_D}_{\mathbf{CRP}}$ can itself be empirically tested and, in Chapter II, we observe that it is almost always satisfied in a laboratory environment; (3) in a stochastic environment this condition is implied by an axiom called *Item Acyclicity*[15] which characterizes an important subset of the general family of Additive Perturbed Utility Models (Fudenberg et al., 2015).

We denote this method as **CRP** $\in \mathcal{G}$. It is then defined as follows:

$$xR^{C_D}_{\mathbf{CRP}}y \text{ if and only if } C_{xy} \geq C_{yx}$$

[14]If P is cyclic the welfare relation has no maximal elements; therefore it is hard to consider the counting revealed preference as a good welfare method.

[15]Consider a stochastic choice rule p as a mapping that assigns a measure $p(A) \in \Delta(A)$ to each menu $A \in \mathcal{X}$. Let $x \succ y$ if $p(x,A) > p(y,A)$ for some $A \ni x,y$ and $x \sim y$ if $p(x,A) = p(y,A)$. A stochastic choice rule p satisfies Item Acyclicity if there exists no sequence $(x_1,\ldots,x_m)$ such that:

$$x_1 \succeq x_2 \succeq \cdots \succeq x_m \succ x_1$$

Sequential

The sequential solution **SEQ** $\in \mathcal{F}$ has been characterized by Horan & Sprumont (2016). This method behaves straightforwardly as a function $\mathbf{SEQ} : \mathscr{C}(\mathcal{X}) \rightarrow \mathcal{B}(X)$. It works recursively such that w.l.o.g. $xP^{*C_{\mathcal{X}}}_{\mathbf{SEQ}}y$ for all $y \in X$, if $x = C(S)$ and $S = \text{argmax}_{S \in D} |S|$ (note that since $D = \mathcal{X}$ we have always $S = X$); then $yP^{*C_{\mathcal{X}}}_{\mathbf{SEQ}}z$ for all $z \in X$, with $z \neq x$ if $y = C(S \setminus \{x\})$; then $zP^{*C_{\mathcal{X}}}_{\mathbf{SEQ}}w$ for all $w \neq y,x$ if $z = C(S \setminus \{x,y\})$ and so on.

If $D \subset \mathcal{X}$, given a choice function $C_D \in \mathscr{C}(D)$, the resulting order is defined as follows: take the set of choice functions $\hat{C}(\mathcal{X}) \in \mathscr{C}$ that extend C_D to $\mathcal{X}$ s.t. $C(S) = \hat{C}(S)$ for all $S \in D$. Then, $\mathbf{SEQ}(C_D)$ is the intersection of the orderings $\mathbf{SEQ}(\hat{C}_{\mathcal{X}})$ assigned to these choice functions. Clearly, the resulting order could be incomplete. This extension is suggested by Horan & Sprumont (2016). Along the paper we apply this extension in the following way: consider the same choice function C_D over an incomplete domain $D \subset \mathcal{X}$; then $\mathbf{SEQ}(C_D) = B \subseteq \mathcal{B}(X)$ where $B = \{P^* \in \mathcal{B}(X) : P^* = \mathbf{SEQ}(\hat{C}_{\mathcal{X}})\}$.

Bernheim, Rangel

Bernheim & Rangel (2009) proposed the following method $\mathbf{BR} \in \mathcal{G}$[16]. For all $x,y \in X$ and for all $D \in \mathcal{D}$, $xP^{C_D}_{\mathbf{BR}}y$ if and only for all $S \in D$ s.t. $x,y \in S$ we have $x = C(S)$ for some S and $y \neq C(S)$ for all S. Otherwise, $xI^{C_D}_{\mathbf{BR}}y$. This method always maps into acyclic binary relations if $D = \mathcal{X}$ - (Bernheim & Rangel, 2009)[Theorem 1]. However since our domains admit multiple and missing observations, $P^{C_D}_{\mathbf{BR}}$ could be cyclic. **BR** can be equivalently defined using the counting revealed preference measure as follows: $xP^{D}_{\mathbf{BR}}y$ if and only if $C_{xy} > 0$ and $C_{yx} = 0$. Otherwise, $xI^{D}_{\mathbf{BR}}y$.

Minimum swaps

This method has been proposed by Apesteguia & Ballester (2015) and denoted as $\mathbf{MS} \in \mathcal{F}$. For all domains $D \in \mathcal{D}$, $\mathbf{MS}(C)$ is defined as follows:

[16] More extensively, they say x is *strictly unambiguously* chosen over y (denote as $x\hat{P}y$) if and only if for all $S \in D$ s.t. $x,y \in S$; $y \neq C(S)$. This asymmetric binary relation can be either considered as itself, that would be the coarsest case, or it could be completed by a symmetric component $\hat{R}$ where $x\hat{R}y$ if and only if $\neg y\hat{P}x$, this would be the finest case. We are going to refer to this latter case.

$$\mathbf{MS}(C) = \underset{P^* \in \mathcal{B}(X)}{\operatorname{argmin}} d_s(C, P^*)$$

where

$$d_s(C, P^*) = \sum_{S \in D} |\{x \in S : xP^*C(S)\}|$$

Transitive core

This method has been proposed by Nishimura (2017). Here, we introduce a variation of his proposal, that is a mapping from complete and reflexive binary relations to transitive but possibly incomplete binary relations. The author wrote: "If she chooses one alternative on some occasions and another on others, then we reveal indifference between these two alternatives" (Nishimura, 2017). This approach is totally in line with Bernheim & Rangel (2009); hence it won't bring novelties with respect to Informational Responsiveness. For this reason, it seems of more interest to found his approach on the **CRP** method (complete and reflexive). Furthermore, given its construction, it only makes sense if the completeness axiom is discarded in the codomain. Consequently, the transitive core method, denoted as $\mathbf{TC} \in \mathcal{T}$, is defined for all domain $D \in \mathcal{D}$ and $x, y \in X$ as follows:

$$xR^C_{\mathbf{TC}}y \quad \Leftrightarrow \quad \begin{cases} zR^C_{\mathbf{CRP}}x \Rightarrow zR^C_{\mathbf{CRP}}y \\ yR^C_{\mathbf{CRP}}z \Rightarrow xR^C_{\mathbf{CRP}}z \end{cases} \quad \forall \quad z \in X$$

Eigenvector centrality

This method exploits the definition of centrality in networks in order to define an order of alternatives. The graph is constructed using the **CRP** method.

The adiacency matrix $A = (\omega_{xy})_{x,y \in X}$ is defined as follows:

$$\omega_{xy} = \begin{cases} C_{xy} \quad \mathbf{if} \quad C_{xy} > 0 \\ \varepsilon \quad \mathbf{if} \quad C_{xy} = 0 \end{cases}$$

with small $\varepsilon > 0$. The elements of the main diagonal are all equal to zero. The eigenvector centrality of $x \in X$, denoted as c^e_x, is:

$$c_x^e = \frac{1}{\lambda_{max}} \sum_{y \in X} \omega_{xy} c_y^e$$

where λ_{max} is the greatest eigenvalue of the adjacency matrix. Perron-Frobenius theorem guarantees that c_x^e is a positive real number.[17][18] Hence, the method $\mathbf{EIG} \in \mathcal{G}$ is so defined: for all $D \in \mathcal{D}$ and for all $x, y \in X$ we have $xR^{C_D}_{\mathbf{EIG}}y$ if and only if $c_x^e \geq c_y^e$.

1.4.1 Counting Revealed preference and Minimum Swaps

The connection between **CRP** and **MS** has been already noted by Apesteguia & Ballester (2015).[19]

We show that if $P_{\mathbf{CRP}}$ satisfies acyclicity then the transitive closure of $P_{\mathbf{CRP}}$ is equivalent to the asymmetric part of the minimum swaps relation $P_{\mathbf{MS}}$. The argument exploits the equivalence between minimizing the number of swaps over all the sets and maximizing the sum of $C_{xy} - C_{yx}$ over all the elements given a strict total order $P^* \in \mathcal{B}(X)$. First, it is straightforward to see that a sum over sets is equivalent to a sum over elements.

Lemma 1. $d_s(C, P^*) = \sum_{S \in D} |\{x \in S : xP^*C(S)\}| = \sum_{x,y \in X} |\{S \in K : y = C(S), x \in S,\ xP^*y\}|$

Proof. Trivial. □

By Lemma 1, the number of swaps can be rewritten as:

$$\sum_{x,y \in X} C_{yx} \quad \text{when} \quad xPy$$

In general the maximum number of swaps is:

[17]By Perron-Frobenius theorem, the constructed revealed preference digraph gives satisfactory results only if it is strongly connected. If it is not (i.e. an element is never chosen); one needs to define a $\varepsilon > 0$ s.t. the eigenvector associated with λ_{max} has strictly positive and real components (it happens only if the adjacency matrix is irreducible; and it is irreducible if the associated digraph is strongly connected).

[18]$\varepsilon > 0$ can be also interpreted as a degree of importance of those elements that are never chosen. If $\varepsilon = 0$ this would mean that beating those elements is worthless. Suppose, for instance, $x = C(x, y, z)$ and $y = C(x, y)$, intuitively $xP^{C_D}_{\mathbf{EIG}}y$ since x beat z. However, if $\varepsilon \to 0$ then the ranking between x, y tends to $xI^{C_D}_{\mathbf{EIG}}y$ because z becomes more and more irrelevant. Actually, if the digraph d_1 is already strongly connected, given the same digraph d_2 but with $\varepsilon > 0$ we have that the $\lim_{\varepsilon \to 0} c_x^e(d_2) = c_x^e(d_1)$.

[19]Apesteguia & Ballester (2015) introduced the following property: A collection of observations satisfies P-Monotonicity if xPy implies $C_{xy} > C_{yx}$. They then established the following result:

Theorem. *If a collection of observations satisfies P-Monotonicity, then P is the unique minimum swaps preference.*

$$\sum_{x,y \in X} |\{S \in K \mid x = C(S), y \in S\}| + |\{S \in K \mid y = C(S), x \in S\}| = \sum_{x,y \in X} C_{xy} + C_{yx}$$

Let's define a new measure $\Delta(C, P^*)$ that equivalently to the swaps distance defines the degree of similarity between a choice function and an irreflexive order P^*:

$$\Delta(C, P^*) = \sum_{x,y \in X} [C_{xy} - C_{yx}] \quad \text{when} \quad xP^*y$$

Lemma 2. *$d_s(C, P_1^*) \leq d_s(C, P_2^*) \Leftrightarrow \Delta(C, P_1^*) \geq \Delta(C, P_2^*)$ for all $P_1^*, P_2^* \in \mathcal{B}(X)$.*

Proof. The proof is algebraic. Note that, given xP^*y:

$$\sum_{x,y \in X} [C_{xy} + C_{yx}] = \sum_{x,y \in X} [C_{xy} + C_{yx}]$$

$$\underbrace{\sum_{x,y \in X} [C_{xy} - C_{yx}]}_{\Delta(C,P^*)} - \sum_{x,y \in X} C_{xy} = - \underbrace{\sum_{x,y \in X} C_{yx}}_{d_s(C,P^*)}$$

Hence, if $d_s(C, P^*)$ increase by n, then it must be that $\Delta(C, P^*)$ decreases by $2n$. Note that, the result is valid for all $C^D, C_1^D \in \mathscr{C}(D)$ since:

$$\sum_{x,y \in X} C_{xy}^D + C_{yx}^D = \sum_{x,y \in X} C_{1,xy}^D + C_{1,yx}^D$$

□

We now prove the result of Apesteguia & Ballester (2015). In this proof we denote $\hat{P}_{\mathbf{CRP}}$ as the transitive closure of $P_{\mathbf{CRP}}$. Recall that since $\mathbf{MS} \in \mathcal{F}$ and $\mathbf{CRP} \in \mathcal{G}$, the indifferences in $R_{\mathbf{CRP}}$ are broken as described in Section 1.2.

Theorem 1. *If $P_{\mathbf{CRP}}$ is acyclic, then $xP^*_{\mathbf{CRP}}y \Leftrightarrow xP_{\mathbf{MS}}y$.*

Proof. By Lemma 2, we can run the proof showing that $P_{\mathbf{CRP}}$ maximizes $\Delta(C, P^*)$. In particular, note that if $P_{\mathbf{CRP}}$ is acyclic and $xP_{\mathbf{CRP}}zP_{\mathbf{CRP}}y$ and $xI_{\mathbf{CRP}}y$, we have that if $x\hat{P}_{\mathbf{CRP}}y$ then $C_{xy} \geq C_{yx}$ for all $x, y \in X$. Hence, $\hat{P}_{\mathbf{CRP}}$ maximizes $\Delta(C, P^*_{\mathbf{CRP}})$. In fact, suppose $yP_{\mathbf{MS}}x$, then by transitivity of $P_{\mathbf{MS}}$, either $zP_{\mathbf{MS}}x$ or $yP_{\mathbf{MS}}z$. Hence, since $C_{xy} = C_{yx}$, $C_{xz} > C_{zx}$ and $C_{zy} > C_{yz}$, we must have that $\Delta(C, P_{\mathbf{MS}}) < \Delta(C, \hat{P}_{\mathbf{CRP}})$, contradicting the definition of $P_{\mathbf{MS}}$. Note that the result holds for any sequence of z_i. □

1.5 Non-triviality of Informational Responsiveness

Firstly, we recall the definition of trivial axiom adopted here. We define an axiom trivial if all the welfare methods proposed by the literature satisfy it. The previous section allows us to prove the *non-triviality* of Informational Responsiveness. The following examples show that neither **BR** nor **SEQ** satisfy Informational Responsiveness.

Example 2. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y,z\} & \{x,y,w\} & \{x,y\} \\ C(S) & (0,1,0) & (1,0,0) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|cc} S & \{x,y,z\} & \{x,y\} \\ C(S) & (0,1,0) & (1,0) \end{array}\right]$$

The two dataset differs in only one observation $x = C(x,y,w)$, however in both cases $xI^{C_D}_{\mathbf{BR}}y$, suggesting that the observation $x = C(x,y,w)$ does not produce any information.

Example 3. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y,z,w\} & \{x,y,z\} & \{x,y,w\} \\ C(S) & (0,0,0,1) & (0,0,1) & (1,0,0) \end{array}\right]$$

$$\left[\begin{array}{c|cc} S & \{x,y,z,w\} & \{x,y,z\} \\ C(S) & (0,0,0,1) & (0,0,1) \end{array}\right]$$

The two dataset differs in only one observation $x = C(x,y,w)$, however in both cases $xP^{*C_D}_{\mathbf{SEQ}}y$ and $yP^{*C_D}_{\mathbf{SEQ}}x$.

These two examples show that both **BR** and **SEQ** fail to satisfy the requirements for Proposition 1 and 2. Therefore, they fail to infer the underlying utility of i.i.d.

RUMs even when it is observed on a single set. The reader may verify that both methods satisfy Choice non-negativeness and Neutrality (trivial axioms); but not Informational Responsiveness. All the other methods introduced, on the contrary, satisfy Informational Responsiveness. The verification of this claim is left to the reader. It is nonetheless instructive to prove it in the case of **TC**.

Claim 1. ***TC** satisfies Informational Responsiveness.*

Proof. First of all we show that $xP^{C_D}_{\mathbf{CRP}}y$ implies either (1) $xP^{C_D}_{\mathbf{TC}}y$ or (2) $\neg xR^{C_D}_{\mathbf{TC}}y$ and $\neg yR^{C_D}_{\mathbf{TC}}x$. (1) is by construction: suppose for all $z \neq x$ the definition is satisfied, take $z = x$ then $xR^{C_D}_{\mathbf{CRP}}x$ by reflexivity and $xP^{C_D}_{\mathbf{CRP}}y$ by assumption; if $z = y$ it is immediate that $\neg yR^{C_D}_{\mathbf{CRP}}x$, hence $xP^{C_D}_{\mathbf{TC}}y$. (2) can be constructed as follows: suppose $xI^{C_D}_{\mathbf{CRP}}z$ and $yP^{C_D}_{\mathbf{CRP}}z$ then by definition $\neg xR^{C_D}_{\mathbf{TC}}y$ and $\neg yR^{C_D}_{\mathbf{TC}}x$ (note that this argument follows also by Axiom 1, called Prudence, of Nishimura (2017)).

Consequently, $xI^{C_D}_{\mathbf{TC}}y$ implies $xI^{C_D}_{\mathbf{CRP}}$. The converse is true only if the definition is satisfied for all $z \in X$. But then, if we add one observation where $x = C(S)$ and $y \in S$ we are in case (1). Thus, Informational Responsiveness is satisfied.

□

A different, stronger, and more naive version of Informational Responsiveness is not satisfied by **MS** and **TC**. We require that if a method ranks x indifferent to y, then an observation of x chosen, even without y available, should make the method rank x better than y.

Axiom 6 (*Strong Informational Responsiveness*)**.**

$$\text{If } xI^{C_D}y \text{ and } x = C(S) \text{ then } xP^{C_{D\cup S}}y \text{ and } yP^{C_{D\setminus S}}x.$$

The following two examples show that both **MS** and **TC** fail to satisfy Strong Informational Responsiveness. In Chapter II, the methods are analysed empirically and the data show that this property is too strong and methods that satisfy it, such as **CC** or **EIG**, are outperformed by those that satisfy only Informational Responsiveness. This argument recalls the importance of standard revealed preference as presented in the introductory section.

Example 4. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} \\ C(S) & (1,0,0) & (0,1) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|cc} S & \{x,y,z\} & \{x,y\} \\ C(S) & (1,0,0) & (1,0) \end{array}\right]$$

The two datasets differ from the observation $x = C(x,z)$. From the first dataset: $xP^{*C_D}_{\boldsymbol{MS}}zP^{*C_D}_{\boldsymbol{MS}}y$ and $xP^{*C_D}_{\boldsymbol{MS}}yP^{*C_D}_{\boldsymbol{MS}}z$ and $yP^{*C_D}_{\boldsymbol{MS}}xP^{*C_D}_{\boldsymbol{MS}}z$. Hence, $xP^{*C_D}_{\boldsymbol{MS}}y$ and $yP^{*C_D}_{\boldsymbol{MS}}x$. However, from the second dataset also $xP^{C_{D\setminus\{x,z\}}}_{\boldsymbol{MS}}y$ and $yP^{*C_{D\setminus\{x,z\}}}_{\boldsymbol{MS}}x$. Hence, Strong Informational Responsiveness is violated.

Example 5. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y\} & \{y,z\} & \{x,z\} \\ C(S) & (3,2) & (1,1) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|ccc} S & \{x,y\} & \{y,z\} & \{x,z\} \\ C(S) & (3,1) & (1,1) & (1,0) \end{array}\right]$$

The two datasets differ from the observation $y = C(x,y)$. From the first dataset $yI^{C_D}_{\boldsymbol{TC}}z$ since $xP^{C_D}_{\boldsymbol{CRP}}y$, $xP^{C_D}_{\boldsymbol{CRP}}z$ and $yI^{C_D}_{\boldsymbol{CRP}}z$. From the second dataset, the same welfare relations hold. Hence, Strong Informational Responsiveness is violated.

1.6 Sensitivity of methods

So far, the proposed conditions have constrained the methods in an informational way. In this section, we focus on a different feature of methods: *sensitivity*. We define two properties that bound the capacity of one observation to influence the welfare relation.

Axiom 7 (*Stability*)**.**

$$\text{If } xP^{C_D}y \text{ then } \neg yP^{C_{D\setminus S}}x \text{ for all } S \in D.$$

Axiom 8 (*Robustness*)**.**

If there exists a z s.t. $xP^{C_D}zP^{C_D}y$ *then* $xP^{C_{D\setminus S}}y$ *for all* $S \in D$.

1.6.1 Normative Interpretation of the axioms

Stability deals with the excessive sensitivity of the method to choice observations around the indifference classes. The stated version, limited to a single observation, is the strongest possible in the context of choice. It asserts that a single choice of y from a set $S \in D$ cannot reverse the judgement from $xP^{C_D}y$ to $yP^{C_{D\cup S}}x$. One can, alternatively, propose weaker versions where the judgement is allowed to be reversed only if the choice comes from sets that are considered particularly "important". However, the reader may note that the level of abstraction limits the definition of "importance" either to the cardinality of sets or to the element chosen.

Both **BR** and **SEQ** methods are stable. Instead both **TC** and **MS** are not stable (the latter example is valid also in the case of **EIG**).

Example 6. *The following two datasets are observed:*

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (1,1) & (0,1) & (1,1) \end{array}\right]$$

$$\left[\begin{array}{c|ccc} S & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,1) & (0,1) & (1,1) \end{array}\right]$$

From the first dataset we have $C_{xy} > C_{yx}$, $C_{yz} = C_{zy}$ *and* $C_{xz} = C_{zx}$. *Applying the definition of transitive core method we obtain* $xP_{\mathbf{TC}}^{C_{D\cup S}}zP_{\mathbf{TC}}^{C_{D\cup S}}y$. *From the second dataset we have* $C_{xy} = C_{yx}$, $C_{yz} = C_{zy}$ *and* $C_{xz} < C_{zx}$, *and by the same principle* $zP_{\mathbf{TC}}^{C_D}yP_{\mathbf{TC}}^{C_D}x$. *Hence, Stability is violated since the two datasets differs from one observation* $x = C(x,y,z)$.

Example 7. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y,z\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (0,1) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|cc} S & \{x,z\} & \{y,z\} \\ C(S) & (0,1) & (1,0) \end{array}\right]$$

From the first dataset: $xP^{*C_D}_{\mathbf{MS}}yP^{*C_D}_{\mathbf{MS}}z$*. However, the same method infers* $yP^{*C_D}_{\mathbf{MS}}zP^{*C_D}_{\mathbf{MS}}x$ *from the second dataset, hence Stability is violated.*

Robustness reproduces the idea that if the researcher is "strongly" convinced that x is better than y than a single choice observation cannot turn her judgement into x indifferent to y. The version of this axiom previously stated translates the idea of x being "strongly" better than y into the statement: there exists a $z \neq x,y$ such that $xP^{C_D}zP^{C_D}y$. In the characterization theorems of the last section, this axiom will be shown to be redundant. However, its interpretation remains relevant because it asserts a normatively important property of the method that will be analysed more in detail in the following subsection. As for the previous axioms, the version proposed can be weakened allowing more alternatives $(z_i)_{i=1}^{n}$ to be between x and y to judge x "strongly" better than y.

Examples 5 and 6 show that **TC**, **MS** and **EIG** violate Robustness. Examples 7 and 8 show that also **BR** and **SEQ** are not robust.

Example 8. *The following two datasets are observed:*

$$\left[\begin{array}{c|ccc} S & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0) & (1,0) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|cc} S & \{x,y\} & \{y,z\} \\ C(S) & (1,0) & (1,0) \end{array}\right]$$

From the first dataset: $xP^{*C_D}_{\mathbf{BR}}yP^{*C_D}_{\mathbf{BR}}z$ *and* $xP^{C_D}_{\mathbf{BR}}z$*. However, from the second dataset:* $xI^{C_D}_{\mathbf{BR}}z$ *violating Robustness.*

Example 9. *The following two datasets are observed:*

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (1,0) & (1,0) & (1,0) \end{array}\right]$$

$$\left[\begin{array}{c|ccc} S & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0) & (1,0) & (1,0) \end{array}\right]$$

From the first dataset the sequential method infers $xP^{*C_D}_{\mathbf{SEQ}}yP^{*C_D}_{\mathbf{SEQ}}z$. *However, from the second dataset:* $xP^{*C_D}_{\mathbf{SEQ}}z$ *and* $zP^{*C_D}_{\mathbf{SEQ}}$ *violating Robustness.*

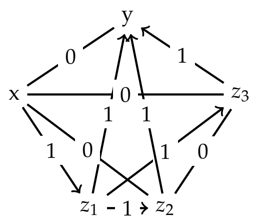
The violation of Robustness by **BR** is of crucial importance. The reader may note that **BR** violates Robustness irrespectively of the number of elements z in between x and y and even more strongly, irrespectively of the number of times x and y have been chosen. In fact, since $xP_{\mathbf{BR}}y$ if and only if y is never chosen, we have that $xI_{\mathbf{BR}}y$ almost surely on the set of all choice functions $\mathscr{C}$. As mentioned in the Introduction, this result shows the excessive cost imposed by the acyclicity of revealed preference relation. In the next subsection, we discuss the topic of the violation of Robustness more in details. We show that **MS**, but also **EIG**, are also highly not robust, but crucially the violation does not hold almost surely as the reader will note from the construction of the counterexample in Example 10.

1.6.2 (Weak) Robustness, large sets and choice overload

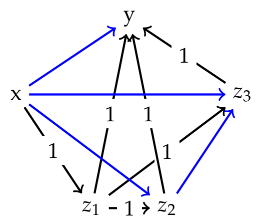
As mentioned in the previous section, one can think of a family of weaker versions of Robustness. In particular, define ϵ as a measure of robustness; then define a sequence $(z_i)_{i=1}^{\epsilon}$ for some $\epsilon < |X| - 2$: if $xPz_1P\ldots Pz_\epsilon Py$ then $xP^{C_{D\setminus S}}y$ for all $S \in D$. In words, less robust axioms require more elements between x and y to guarantee that the asymmetric relation is preserved. In the introduction, we described Stability and Robustness as continuity requirements. The reader may note how this definition resembles the one of $\varepsilon - \delta$ continuity. We extensively discuss this intuition in Appendix A.2.

This new definition allows us to measure the degree of Robustness of welfare methods. Here, we propose an example that is illuminating with regard to the minimum swaps welfare relation. That is, for any ε, **MS** is not robust. The origin of this result is the excessive weight assigned by **CRP** to large sets. The following example, which exploits the result of Theorem 1, not only shows that Robustness is violated, but even more strongly that even Stability is violated.

Example 10. *Let* $X = \{x, y, z_1, z_2, z_3\}$. *The first graph describes* ***CRP*** *on some choice function* C. *The second graph is its transitive closure.*

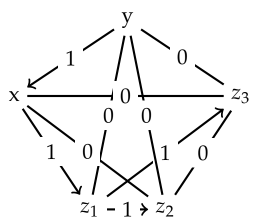


$xP^C_{\mathbf{CRP}}z_1P^C_{\mathbf{CRP}}z_2$, $z_1P^C_{\mathbf{CRP}}z_3P^C_{\mathbf{CRP}}y$

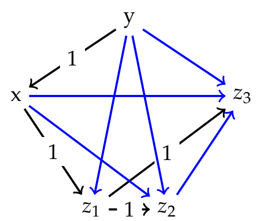


$xP^C_{\mathbf{MS}}z_1P^C_{\mathbf{MS}}z_2P^C_{\mathbf{MS}}z_3P^C_{\mathbf{MS}}y$

Suppose that following observation $y = C(x, z_1, z_2, z_3, y)$ *is added to the dataset:*



$yP^C_{\mathbf{CRP}}xP^C_{\mathbf{CRP}}z_1P^C_{\mathbf{CRP}}z_2$, $z_1P^C_{\mathbf{CRP}}z_3$



$yP^C_{\mathbf{MS}}xP^C_{\mathbf{MS}}z_1P^C_{\mathbf{MS}}z_2P^C_{\mathbf{MS}}z_3$

A comment has to be made concerning an intrinsic problem connected with revealed preference. Methods based on revealed preference tend, by construction, to attribute a larger weight to observations from larger sets. However, this tendency seems to be in contrast with the literature of choice overload.[20] For instance, if one assumes that individuals do more mistakes in large sets, then he can think that the inference of welfare should be less influenced by such observations. We will empirically test this hypothesis in Chapter II.

1.7 Counting Procedures

In this section, we show how the introduced axioms play a role in characterizing the two counting procedures **CC** and **CRP**.

[20]Examples of models that produce more inconsistencies in bigger sets are Frick (2016) and Fudenberg et al. (2015).

1.7.1 Counting choice method

Firstly, we characterize the *counting choice method* [**CC**]. This characterization has some similarities with the one proposed by Rubinstein (1980) for scores in tournaments and van den Brink & Gilles (2003) for outdegrees of digraphs. However, we deal with the higher complexity of a domain of choice menus. For this reason, the axioms proposed are stronger. Strong Informational Responsiveness and Stability imply Strong Positive Responsiveness (axiom used by both Rubinstein (1980) and van den Brink & Gilles (2003)); and this latter is shown to be not sufficient to characterize **CC**. We refer to Strong Positive Responsiveness as defined in Section 1.2. in the footnote to the definition of Informational Responsiveness.

Theorem 2. *A method g satisfies Stability, Independence, Strong Informational Responsiveness and Neutrality if and only if $g =$ **CC**.*

Proof. The *only if* part is trivial. We prove only the *if* part.

Step 1 (Induction base).

The proof is by induction over the cardinality (number of non-empty sets in the domain) of the domain D given a generic choice function C_D. Let's prove the statement for $|D| = 2$. Suppose $D = \emptyset$, we have $xI^{C_{D_\emptyset}}y$. In fact, suppose $xP^{C_{D_\emptyset}}y$ then by Neutrality $yP^{\pi(C_{D_\emptyset})}x$ if $\pi(x) = y$ and $\pi(y) = x$; but $g : \mathcal{C} \to \mathcal{R}(X)$ and we would have two orders R_1, R_2 associated with the same choice function s.t. xP_1y and yP_2x, hence by Completeness $xI^{C_{D_\emptyset}}y$.[21] Suppose $D = \{S\}$. If $z = C(S)$ then by Independence $xI^{C_D}y$ and clearly $C_x = C_y$. If $x = C(S)$ then by Strong Informational Responsiveness $xP^{C_D}y$ and $C_x = 1$, $C_y = 0$. Note that $|D| = 1$ is not enough for our purpose. In particular, we have to prove also that $C_x = C_y = 1 \Leftrightarrow xI^{C_D}y$ with $|D| = 2$ since we need x,y chosen in the domain to make the base general over all the possible domains.[22] So, suppose $D = \{S,T\}$. If $z = C(T)$ the result holds by

[21] If we rewrite the subsequent of Stability as $xR^{C_D}y$, instead of $\neg yP^{C_D}x$; then the role of Neutrality and Completeness is restricted to implying: $D = \emptyset \Rightarrow xI^{C_{D_\emptyset}}y$ for all $x,y \in X$. One can eventually assume this condition as axiom and prove the theorem without using these two axioms.

[22] In order to see why this last part is necessary, let's suppose to prove the base for induction using Strong Positive Responsiveness. Clearly, if $|D| = 1$ the statement is proved to be true. Suppose we take a domain $|D| = n$ where the statement is true and $|D \cup \{T\}| = n + 1$. We can prove that $C_x > C_y \Rightarrow xP^{C_D}y$. However, Example 7 provides a method, different from the counting choice method, that satisfy Strong Positive Responsiveness, Neutrality and Independence but not the last statement. This

Independence. If $x = C(S) = C(T)$ then by Stability we have $\neg y P^{C_D} x$. Suppose, by Completeness, $x I^{C_D} y$ then by Strong Informational Responsiveness we should have $y P^{C_{D \setminus T}} x$ contradicting the result at $|D| = 1$; hence $x P^{C_D} y$ with $C_x = 2$ and $C_y = 0$. If $C(S) = x$ and $C(T) = y$ then we have $y P^{C_{D \setminus S}} x$ and by Stability $\neg x P^{C_D} y$ and $x P^{C_{D \setminus T}} y$ and by Stability $\neg y P^{C_D} x$; hence by Completeness $x I^{C_D} y$ with $C_x = C_y = 1$.

Step 2 ($C_x \geq C_y \Rightarrow x R^C y$).

Suppose $|D| = n$ and the statement holds. Take then $|D \cup \{T\}| = n + 1$. Suppose that $C_x = C_y$. If $z = C(T)$, $x I^{C_{D \cup T}} y$ by Independence and the inductive hypothesis. Suppose $x = C(T)$ then if we take out T by inductive hypothesis we have $y P^{C_D} x$ and by Stability $\neg x P^{C_{D \cup T}} y$. But then, since $C_x = C_y$ there exists a set S such that $y = C(S)$. Hence by inductive hypothesis $x P^{C_{D \cup T \setminus S}} y$ and so $\neg y P^{C_{D \cup T}} x$ by Stability, which means by Completeness $x I^{C_{D \cup T}} y$.

Suppose $C_x > C_y$. If $z = C(T)$, $x P^{C_{D \cup T}} y$ by Independence and the inductive hypothesis. If $x = C(T)$ then if we take out T, we have two scenarios: if $x I^{C_D} y$ then $x P^{C_{D \cup T}} y$ by Strong Informational Responsiveness (note that this is the case when $C_x - C_y = 1$). If $x P^{C_D} y$ (when $C_x - C_y > 1$) then $\neg y P^{C_{D \cup T}} x$ by Stability. However, suppose by contradiction $x I^{C_{D \cup T}} y$, then taking out $x = C(T)$ we should have $y P^{C_D} x$ contradicting the inductive hypothesis, so by Completeness $x P^{C_{D \cup T}} y$. If $y = C(T)$ then if we take out T, by inductive hypothesis $x P^{C_D} y$ and by Stability $\neg y P^{C_{D \cup T}} x$. However, since $C_x > C_y$ there exists a set S s.t. $x = C(S)$ and so by the previous argument $x P^{C_{D \cup T}} y$.

□

Claim 2. *Strong Informational Responsiveness and Stability ⇒ Strong Positive Responsiveness.*

Proof. Strong Informational Responsiveness proves Strong Positive Responsiveness when the antecedent is $x I^{C_D} y$. Suppose $x P^{C_D} y$; if we add $x = C(S)$ then by Stability it cannot be $y P^{C_{D \cup S}} x$. Suppose, by Completeness $x I^{C_{D \cup S}} y$, then by Strong Informational Responsiveness we have $y P^{C_D} x$ which contradicts the initial condition $x P^{C_D} y$.

suggest that the strategy of the proof is fallacy in some parts. In fact, if $C_x = C_y$ and we focus on $|D| = 1$ we do not cover all the choice functions where x or y are chosen. Hence, since the base must have a universal quantifier, the proof would be incomplete.

□

The following examples provide two methods that satisfy Independence, Neutrality, Strong Positive Responsiveness ($N_{xy} \geq N_{yx}$) and Transitivity ($Q_x \geq Q_y$) but are not the counting choice method.

Example 11. *Given two elements $x, y \in X$ and a choice function $C_D \in \mathscr{C}(D)$:*

$$N_{xy} \geq N_{yx} \Leftrightarrow x R^{C_D} y$$

where $N_{xy} = C_{xy} + \delta \cdot |\{S : x = C(S), y \notin S\}|$ with $\delta \in (0,1)$.

Example 12. *Given an element $x \in X$ and a choice function $C_D \in \mathscr{C}(D)$:*

$$Q_x \geq Q_y \Leftrightarrow x R^{C_D} y$$

where $Q_x = \sum\limits_{S : x = C(S)} |S|$.

However, if we restrict the domain on binary sets then Strong Positive Responsiveness and Transitivity become sufficient.

Proposition 3 (Theorem 1 - Rubinstein (1980))**.** *Let D be a domain of solely binary sets. A method g satisfies Strong Positive Responsiveness, Independence, Neutrality and Transitivity if and only if $g = \boldsymbol{CC}$.*

Proof. See Appendix A.1. □

Independence of the axioms

Stability

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (1,0) & (0,2) & (2,0) \end{array}\right]$$

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (1,0) & (0,1) & (2,0) \end{array}\right]$$

Following the method proposed in Example 10 and setting $\delta = 0.5$ we obtain: $N_{xy} = 2 > N_{yx} = 1$; $N_{xz} = 1.5 < N_{zx} = 2$; $N_{yz} = 2 > N_{zy} = 1$. Hence, $x P_g^{C_D} y P_g^{C_D} z P_g^{C_D} x$

(Transitivity is violated). From the second dataset $N_{xy} = 2 > N_{yx} = 1$; $N_{xz} = 1.5 > N_{zx} = 1$; $N_{yz} = 2 > N_{zy} = 0.5$, hence $xP_g^{C_D}yP_g^{C_D}z$, $xP_g^{C_D}z$ violating Stability since the two datasets differ only from one observation $z = C(x,z)$.

Strong Informational Responsiveness

CRP satisfies Stability, Neutrality, Independence but only the weaker version of Informational Responsiveness.

$$C_{xy} \geq C_{yx} \Leftrightarrow xR_{\mathbf{CRP}}^{C_D}y$$

Independence

Let $T_{xy} = |\{S \in D : x \in T, y \notin T, x \neq C(T)\}|$. The following method g is defined:

$$C_x + T_{xy} - C_y - T_{yx} \geq 0 \Leftrightarrow xR_g^{C_D}y$$

Note that this method satisfies Stability because a single observation from a set $S \in D$ can increase the score by maximum one. If $x = C(S)$ then $|T_{xy}|$,$|T_{yx}|$ and C_y don't change. Similarly if $y = C(S)$. If $z = C(S)$ and either $x,y \in S$ or $x,y \notin S$ the order between x,y is not affected. If $x \in S$ and $y \notin S$ then only $|T_{xy}|$ increases. However, as the following example shows, Independence is violated:

$$\left[\begin{array}{c|cccc} S & \{x,y,z\} & \{x,y\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (1,0) & (0,2) & (2,0) \end{array}\right]$$

$$\left[\begin{array}{c|ccc} S & \{x,y,z\} & \{x,z\} & \{y,z\} \\ C(S) & (1,0,0) & (0,2) & (2,0) \end{array}\right]$$

From the first dataset we infer $xP_g^{C_D}yP_g^{C_D}zP_g^{C_D}x$, while from the second dataset $zP_g^{C_D}xP_g^{C_D}y$ and $yI_g^{C_D}z$. Hence, the observation $x = C(x,y)$ has modified the welfare relation between y,z violating Independence.

Neutrality

$$\begin{cases} \forall x, z \neq y \quad \Rightarrow \quad [C_x \geq C_z \Leftrightarrow x R_g^{C_D} z] \\ \forall x \quad \Rightarrow \quad x P_g^{C_D} y \end{cases}$$

1.7.2 Counting Revealed Preference method

Additional axiom

It is straightforward to see that the **CRP** method doesn't satisfy Strong Informational Responsiveness. However, it satisfies Informational Responsiveness. But, this is not the only difference between the two. If it was, then there would be an inclusion relation between the methods, which is not the case. Hence, the following axiom is introduced.

Axiom 9 (*Connection*)**.**

$$\textit{For all } S \in D \textit{ s.t. } \{x, y\} \nsubseteq S \textit{ then } x R^{C_D} y \Leftrightarrow x R^{C_{D \setminus S}} y.$$

The interpretation of this axiom is quite clear. Intuitively it makes Informational Responsiveness much stronger. In fact, together, they don't only require that each set S s.t. $x = C(S)$ and $y \in S$ produce some information about x, y; but they require that these are the only sets doing that.

Theorem 3. *g satisfies Neutrality, Stability, Informational Responsiveness and Connection if and only if $g =$ **CRP**.*

Proof. See Appendix A.1. □

The reader may note that Independence is implied by the other axioms and so redundant. It is clearly not true that Connection and Independence are equivalent under the other axioms. As shown in the subsection about the independence of the axioms, **CC** satisfies Neutrality, Informational Responsiveness, Stability, Independence but not Connection.

Corollary 2. *Neutrality, Connection, Informational Responsiveness and Stability imply Independence.*

Remark 1. Note that **CC** and **CRP** are not nested. In particular, **CC** satisfies Strong Informational Responsiveness while **CRP** does not. Conversely, **CRP** satisfies Connection while **CC** does not. Nonetheless, both satisfy Neutrality, Stability, Independence and Informational Responsiveness.

Independence of the axioms

The reasoning behind these examples is the same discussed for**CC**.

Stability

Similarly to Example 10, define the following method:

$$Q_{xy} \geq Q_{yx} \Leftrightarrow xR^{C_D}y$$

where $Q_{xy} = \sum\limits_{S:x=C(S),y\in S} |S|$.

This method does not satisfy Stability since the value attached to the sets depends on their cardinality:

$$\left[\begin{array}{c|cc} S & \{x,y,z,w,t\} & \{x,y\} \\ C(S) & (1,0,0) & (0,2) \end{array}\right]$$

$$\left[\begin{array}{c|c} S & \{x,y\} \\ C(S) & (0,2) \end{array}\right]$$

From the first dataset $Q_{xy} = 5 > Q_{yx} = 4$ while from the second dataset $Q_{xy} = 0 < Q_{yx} = 4$.

Informational Responsiveness

The following method satisfies Connection, Neutrality and in a vacuous way also Stability. However, it violates Informational Responsiveness.

$$xI^{C_D}y \quad \forall \quad x,y \in X$$

Connection

CC satisfies Neutrality, Stability and (Strong) Informational Responsiveness, however it doesn't satisfy Connection.

$$C_x \geq C_y \Leftrightarrow xR^{C_D}y$$

Neutrality

$$\begin{cases} \forall x, z \neq y \quad \Rightarrow [C_{xz} \geq C_{zx} \Leftrightarrow xR^{C_D}z] \\ \forall x \quad \Rightarrow \quad xP^{C_D}y \end{cases}$$

1.8 Summary

Figure 1.1 summarizes the characteristics of the methods analysed in Chapter I.

It is important to notice that for reasons of simplicity of exposure, especially in view of Chapter II, we substitute incompleteness with indifference. These process, that allows a consistent comparison across methods, can undermine the theoretical foundations of some of these methods. Particularly, **MS** and **TC** are affected; although differently. Both methods satisfy IR. However, **MS** satisfies it even when indifferences are introduced; while **TC** does not. Therefore, we treat **MS** with indifferences and **TC** with incompleteness. Hence, **TC** satisfies both transitivity [T] and quasi-transitivity [QT];[23] while **MS** satisfies only QT.

	NEU	CNN	IR	SIR	IND	ST	ROB	CON	QT	T
CRP	$\surd$	$\surd$	$\surd$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\times$	$\times$
MS	$\surd$	$\surd$	$\surd$	$\times$	$\times$	$\times$	$\times$	$\times$	$\surd$	$\times$
TC	$\surd$	$\surd$	$\surd$	$\times$	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$
EIG	$\surd$	$\surd$	$\surd$	$\surd$	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$
CC	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\times$	$\surd$	$\surd$
SEQ	$\surd$	$\surd$	$\times$	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$	$\surd$
BR	$\surd$	$\surd$	$\times$	$\times$	$\surd$	$\surd$	$\times$	$\surd$	$\times$	$\times$

Figure 1.1. **Summary of the properties of welfare methods.**

[23] A binary relation R is quasi-transitivity if the asymmetric part P is transitive (Sen, 1969).

$$\begin{tikzcd}
\mathrm{ST} \wedge \mathrm{SIR} \arrow[r, mapsto] & \mathrm{SPR} \arrow[r, mapsto] \arrow[dr, mapsto] & \mathrm{SIR} \arrow[r, mapsto] & \mathrm{IR} \\
\mathrm{ST} \wedge \mathrm{IR} \wedge \mathrm{CON} \arrow[rr, mapsto] & & \mathrm{CNN} \arrow[r, mapsto] & \mathrm{IR} = \mathrm{IR}^{+}
\end{tikzcd}$$

Figure 1.2. **Implication Diagram "Monotonicity"**

Figure 1.2 presents an implication diagram that deconstruct the property of monotonicity in observations in order to understand its full implications. We denote (Strong) Positive Responsiveness as SPR. This axiom is a strong monotonicity axiom in choices and has been defined at footnote 12. We denote IR as IR^{+} when only adding observations is considered in the definition.

The diagram is divided into two levels. The first is related to **CC**. The simple counting satisfies SPR via ST and SIR. Consequently, it also satisfies CNN. We highlight how CNN guarantees that we can only focus on adding observations to a dataset when we define IR. This result can be found as Claim 1 in Appendix A.1.2. We report a nice counterexample in the absence of CNN. The second level is related to **CRP**. As clear from Theorem 3, **CRP** does not imply SIR, hence also SPR. However, it implies CNN. In this case, it is easy to prove that CNN is redundant in Theorem 3, and that CNN is implied by ST, IR and CON.

1.9 Conclusion

In Chapter I, we analyse the problem of a researcher that wants to elicit the preferences of individuals that have heterogeneous behavioural models. Given the complexity of the task, we propose some simple and normatively appealing properties. Firstly, we show that a property called Informational Responsiveness has important empirical implications since it is crucial to infer the underlying utility of a broad family of stochastic models of choice. Secondly, we propose some "continuity" requirements that constrain the importance of single observations in determining the elicited preference relation. We analyse all the welfare methods proposed by the literature in view of the introduced normative principles. Finally, we completely characterize two counting procedures on datasets with missing data and multiple observations. These procedures will be an important theoretical base for the experimental study that will constitute Chapter II.

Chapter 2

Behavioural Welfare Analysis and Revealed Preference: Experiment

2.1 Introduction

In this Chapter, we test our theoretical contribution using novel experimental data. The researcher's problem here is brought to data. We elicit preferences from a group of 145 subjects and then test if Informational Responsiveness[1] and standard Revealed Preference constitute a solid base for welfare analysis.

All the difficulties of the researcher problem expressed in Chapter I are present here. We study two environments: time and risk preferences. Henceforth we refer to the environments as "Time" and "Risk". We observe a high heterogeneity in behavioural models not only across subjects but also within subjects and across environments. We do not adopt a model-driven approach because even for our simple experimental setting the literature has proposed several models, often mutually exclusive, to explain some of the patterns in the data.[2] Each model provides a different way to construct a so called "revealed" preference relation. Therefore, as discussed in length in the previous chapter, we rely on our simplified model-free approach

[1] We refer to Section 1.3 of Chapter I for an analysis of this axiom.

[2] The existence of different models that explain similar situations regards, for instance, how individuals deal with complex choice problems, in particular when the number of alternatives is high. Both in deterministic and stochastic literature two main lines of models have been developed: (i) (degenerate) attention models has been developed among many by Masatlioglu et al. (2012), Lleras et al. (2017), Manzini & Mariotti (2014), Echenique et al. (2018), Cattaneo et al. (2018); (ii) (uniform) attention models by Frick (2016), Fudenberg et al. (2015).

and test its effectiveness.

We aim particularly to answer the following questions that constitute together the premise and the testing of our theoretical proposals.

- Premise: Do individuals consistently reveal welfare in different choice problems, e.g. in Time and Risk?
- Test: If not, how should a researcher measure welfare when individuals violate the Weak Axiom of Revealed Preference? Particularly: is *Informational Responsiveness* effective in discriminating welfare methods? And how important are revealed preference relations?

To address these questions, we design a new choice elicitation experiment. Subjects are asked to choose from sets that include delayed payment plans (Time) or lotteries (Risk). As in Manzini et al. (2010) we collect choices regarding four alternatives in every subset. Henceforth we refer to them as MAIN alternatives and to the subsets as MAIN sets.[3] The remaining questions contain either problems of asymmetric dominance[4] or choice overload.[5] This structure allows us to test if choices from sets that are potentially doomed by behavioural effects are relevant to elicit preferences.

In order to test the capacity of eliciting preferences, at the end of the experiment, we ask subjects to rank the four MAIN alternatives. We consider this relation as a benchmark for evaluating how welfare methods perform on the dataset. The reliability of the reported preference relation is empirically strong.[6] In an exercise, that we call "Identification", we measure the proportion of subjects for whom each

[3]In Appendix B.1 and B.2 the reader can find descriptions of the alternatives and questions with particular reference to the MAIN ones.

[4]Asymmetric dominance deals with ternary sets where one alternative is clearly dominated by one of the other while the remaining ones have similar value. In this cases subjects typically show attraction effect, e.g. Huber et al. (1982) and Natenzon (2019).

[5]With choice overload we intend a situation where the number of alternatives in a choice set makes it difficult for the decision maker to evaluate all of them. An empirical example can be found in Iyengar & Kamenica (2010).

[6]The reliability of the reported ranking is confirmed by the following statistics: in time preferences 69 out of 70 rational subjects reported the correct optimal alternative and 61 out of 70 reported correctly the entire welfare relation. This statistic is repeated in risk preferences with respectively 10 out of 12 subjects reporting the correct optimal alternative and 9 out of 12 the correct welfare relation. Two subjects reported the opposite ranking to the one they rationally employed in their choices. This probable mistake does not affect our results since every method will clearly fail to identify these subjects.

welfare method, that relies only on choices, can elicit either the entire reported preference relation or simply the reported best alternative.

To the best of our knowledge, this is the first paper that compares elicited preferences with a benchmark. The choice of this benchmark requires some discussions. Previous papers, (Bouacida & Martin, 2020), (Manzini et al., 2010), focused on the properties of the elicited preference relation. However, as discussed in Chapter I, difference methods map into different binary relations, therefore the analysis of the properties of the resulting binary relations is a biased indicator of their efficacy. Another approach may be to measure two similarity measures: (1) among the elicited preference relations across methods with the assumption that methods that report more similar results are more likely to report the true preference; (2) among choice functions, with the assumption that more similar choice functions should be mapped into similar preference relations. However, even these two analysis are debatable. The first one because methods based on similar assumptions mechanically elicit more similar preference relation. The second one because it relies on the choice of the similarity measure on the space of choice functions.

To overcome these difficulties, we decide to use the directly reported preferences as benchmark. This choice is in line with liking-rating tasks as in Reutskaja et al. (2011) where they have been used as measures of values, with the only difference that our liking-rating task is ordinal and not cardinal. Two main issues regard this choice.

First, there may be some misalignments in the reporting between choices and preferences also due to the fact that the latter are not incentivized. Incentives are unlikely to play a role as shown recently by Enke et al. (2021), and are often missing in liking-rating tasks (Reutskaja et al., 2011). However, we control for possible effects. In Section 2.3.5, we discard what Fudenberg et al. (2019) called irreducible error. Namely, those subjects that cannot be identified by any methods because, for instance, they chose according to one preference relation and reported the exact opposite. In this way, our analysis on the performance of methods is constrained on the subjects that show a certain degree of alignment between choices and preferences. Conditional on this set of subjects, the comparison between methods is hardly affected by the incentive mechanism.

Second, the reader may ask why the reported preferences should be considered as the true ones, and why if so, we cannot just elicit preferences asking directly for them. The first concern is well-posed but it relies on the problem of deciding what is the "true" preference, which can be answered only by assumption. In fact, the problem has been generally avoided by the literature. For instance, Bouacida & Martin (2020) evaluate "goodness" of methods using properties such as: number of cycles or completeness of the resulting welfare relation. Here we rely on the assumption that the direct report of the ranking, at the end of the experiment, and when the information about the alternatives have been processed, creates a credible benchmark. Given the tautological nature of the question, we do not argue in favour of its truthfulness, however it allows to overcome the above mentioned biases that characterize the mere comparison of methods. Finally, note that avoiding the question (e.g. focusing on the properties of the inferred preferences) would imply implicit assumptions on the properties of the "true" preferences. The second concern regards the question: "why not asking instead of inferring from choices?". Outside the experimental setting asking for direct reporting of preferences is not an option since often only dataset of choices are available. In an experimental setting, we go back to our previous point on what we believe to be the "true" preference. In Caliari (2020), we investigate the relationship between the characteristics of the elicited and reported preferences and a series of observables such as response times, cognitive abilities, elicited heuristics, etc... In this paper, we open the black box of the decision process and investigate why people prefer certain objects and why eventually they reported different ones. Nonetheless, even though these are fundamental questions, they outside the scope of this Chapter.

Finally, we address questions regarding which choice problem better reveals preferences and what is the connection between consistency and preference revelation. In doing so, we face the well-known problem of comparing consistency of choice among different and non-symmetric parts of the dataset.[7] We solve this problem by developing a measure of consistency that is robust to the structure of

[7]See Andreoni et al. (2013) for a survey of the literature.

the dataset.

2.1.1 Preview of the results

First, we find that a good proportion of subjects never violate WARP in Time (37%). Conversely, and in line with the literature (Agranov & Ortoleva, 2017), almost no subjects satisfy WARP in Risk (6%). The average number of violations of WARP reflects this finding: the average in Time is 11.26 while in Risk is 24.65 (the difference is significant with $p \approx 0$), and robust if we focus only on subjects that violate WARP at least once. In both environments, subjects are not behaving randomly (the average number of violations for a random chooser is 56.70).

Second, we observe that methods that satisfy Informational Responsiveness (IR) outperform the other welfare methods. When asked to uniquely identify the best reported alternative, the Pareto approach (**BR**) is outperformed by 30% in Time and 50% in Risk.[8] When limited to a set identification exercise, more in line with its conservative approach, it is still outperformed by 15% in Time and 20% in Risk. These results are robust when we limit ourselves to the sets that contain only the MAIN alternatives. Similarly, when asked to uniquely identify the entire welfare relation, the Pareto approach is outperformed by 20% in Time and 25% in Risk.

Third, we compare the identification power of the simple counting (**CC**), that satisfy a stronger version of Informational Responsiveness (SIR), with the counting revealed preference procedure (**CRP**). We find that the former is outperformed by 6% in Time and 4% in Risk. This suggests on one hand that IR is not sufficient and that a stronger version could have negative effects; on the other hand, that a notion of frequency in line with standard revealed preference plays an important role in the identification process.

Four, we analyse these results using a measure of completeness for models developed by Fudenberg et al. (2019). The main advantage of this measure is to provide a power of methods with respect to the most naive and most sophisticated method. We use Bernheim & Rangel (2009) approach (**BR**) as most naive method

[8] These percentages are calculated on the total number of subjects. For example, in Time the method proposed by Bernheim & Rangel (2009) uniquely identifies the correct best alternative of 59% of the subjects while the counting revealed preference method of 87%.

and an optimal weighting algorithm (**OW** - Section 2.3.5) as most sophisticated one.[9] The idea is as follows: subjects that are not identified by **OW** are considered an irreducible error; while subjects that are identified by **BR** are considered trivial. We confirm that methods that satisfy IR and are based on standard revealed preference are significantly more complete.

Five, we directly test IR. We combine results on preference elicitation from the optimal weighting algorithm and on consistency of choice from our index of rationality. We find that, in Time, asymmetric dominance particularly increases inconsistency and these sets are the only ones to which the algorithm assigns negative weights. To all other sets, the algorithm associates positive weights confirming the importance of IR.

2.1.2 Related Literature

This Chapter firstly relates to the few choice elicitation experiments such as Manzini et al. (2010) and Barberá & Neme (2017). From those it differs in two main ways: (i) we collect choices on a much richer set of questions to test how behavioural effects affect welfare revelation; (ii) we ask subjects to directly report their preference relation. The experiment by (Manzini et al., 2010) has been analysed from a welfare perspective by Bouacida & Martin (2020), but their analysis is limited to **BR** methods and therefore does not focus on the comparison between different methods. Secondly, our experimental design relates to the literature on stochastic choice and choice deferral. However, even if our design shares some features with existent experiments, none of the following elicit both choices and preferences, is based on both time and risk preferences, and collects choices regarding all non-empty subsets of the MAIN alternatives as well as sets with behavioural effects. Some are restricted to binary comparisons: Agranov & Ortoleva (2017), Hey & Carbone (1995), Danan & Ziegelmeyer (2006), Hey (2001), Cavagnaro & Davis-Stober (2014), Sopher & Narramore (2000), Chabris et al. (2009). Others collect data only on particular sets: Harbarugh et al. (2001) elicited choices from 11 different sets with cardinality from 3 to 7; Iyengar & Kamenica (2010) elicited choices from sets of either 3 or 11 gam-

[9] The optimal weighting algorithm is a data-driven method. We optimally set the weights of the questions in order to maximize the Identification exercise.

bles; Haynes (2009) collected response times but he elicited choices only from sets of either 3 or 10 prizes; Iyengar & Lepper (2000) elicited choices from sets of either 6, 24 or 30 alternatives; Sippel (1997) elicited 10 choices from budget sets regarding 8 alternatives.

The index of rationality (Section 2.3.2) based on the perturbation of a data generating process such as the logit model is connected with the literature on rationality indexes and power measures. The most prominent example is the Selten measure (Selten, 1991). Examples of power tests against random behaviours have been proposed by Becker (1962) and Bronars (1987). Our index is robust to the dataset structure. An example clarifies this statement. Imagine to have a series of datasets of choices ordered by the number of sets involved; bigger is the dataset and higher is the probability of making a mistake or violating WARP. Therefore, if we simply compare the number of mistakes in the different datasets we incur in a clearly biased comparison. Our proposal allows for these comparisons, solving a problem that is common to other indexes such as Afriat's index (Afriat, 1972), minimum number of observations to remove to rationalize the data (Houtman & Maks, 1985), number of violations of consistency axioms (Swofford & Whitney, 1987) and (Famulari, 1995), minimum number of swaps (Apesteguia & Ballester, 2015). A comprehensive review of the literature is offered by Andreoni et al. (2013), and an example of problematic estimates of violations of consistency can be found in Beatty & Crawford (2011).

2.1.3 Experimental Hypothesis

Figure 2.1 reports the characteristics of the welfare methods[10] introduced in Chapter I and constitutes the reference point for the experimental analysis. We test the joint importance of Informational Responsiveness (IR) and Revealed Preference (RP). For this latter, we intend that the foundation of the welfare methods is **CRP**. The reader may note that on one hand two methods, **BR** and **SEQ**, do not satisfy IR with the former based on RP. On the other hand, **CC** satisfies IR but it is not based on RP. In the experimental analysis, we show that these two conditions are both neces-

[10]Descriptions of the requirements and the welfare methods are explained in detail in Chapter I.

sary. Together, they guarantee better solutions to the preference elicitation problem. Furthermore, the variety of methods that satisfy both conditions prove that the researcher has still a high degree of discretion in performing welfare analysis.

The RP assumption requires two comments. Firstly, it contains the trade-off between the very strong ex-ante acyclicity requirement that **BR** imposes on **CRP** and methods that allow to break cycles in **CRP** such as **MS**, **TC** and **EIG**, therefore reestablishing acyclicity ex-post. Secondly, the reader may note that these methods are differently robust. In particular, as shown in Section 1.6, **BR** is infinitely nor robust, while **MS**, **TC** and **EIG**, even if all violate Robustness (and Stability), they do it with different degrees, with **EIG** being the less robust among the three. In Section 2.3.4, we will report a measure of the different degree of continuity of these methods and their consequences on the empirical results.

	NEU	CNN	IR	RP
CRP	$\surd$	$\surd$	$\surd$	$\surd$
MS	$\surd$	$\surd$	$\surd$	$\surd$
TC	$\surd$	$\surd$	$\surd$	$\surd$
EIG	$\surd$	$\surd$	$\surd$	$\surd$
CC	$\surd$	$\surd$	$\surd$	$\times$
SEQ	$\surd$	$\surd$	$\times$	$\times$
BR	$\surd$	$\surd$	$\times$	$\surd$

Figure 2.1. **Properties of welfare methods.**

2.2 Experimental design

The experiment follows a standard choice elicitation design, e.g. Manzini et al. (2010), Barberá & Neme (2017). The complete instructions and screenshots are presented in the Appendix B.4 and B.5. Subjects received instructions both on screen and on paper such that they could consult them during the experiment.

The experiment is divided into three parts: (1) Choice elicitation; (2) Questionnaire; (3) Raven Test. The choice elicitation part has 50 questions; half regarding choice among lotteries (Risk Preference Elicitation) and half regarding choice among delayed payment plans (Time Preference Elicitation). No question was repeated. At the beginning of each part, subjects answered three trial questions in order to make them familiar with the experimental environment.

For both Time and Risk, the alternatives were divided into two groups: four MAIN alternatives, which are presented in Table 2.1 and Table 2.2, and some "confounding" alternatives that are described in Appendix B.1. Each individual solved all the 11 choice problems (6 binary, 4 ternary and the quaternary set) involving the MAIN alternatives, denoted as MAIN sets. The other questions were set in order to obtain particular information about rationality: Monotonicity, Impatience,[11] Stochastic Dominance; and about possible behavioural effects: choice overload, compromise effect, attraction effect. The structure of the questions is presented in Appendix B.2. The positions of the alternatives were randomized. The subjects could face two orders of questions and also we inverted Time and Risk elicitation such that we had a total of four treatments (Appendix B.2.).[12]

One of the fundamental feature of the design is the collection of all the non-empty subsets of the MAIN alternatives with cardinality greater than two. Its importance can be summarized into two motivations: (1) this part of the dataset is symmetric and therefore allows for immediate comparisons across alternatives and environments (Time and Risk). Such comparisons are not straightforward as it will be clear in Section 2.3.3.[13] (2) beyond the symmetry, that is also satisfied by the collection of binary sets, the MAIN sets allow to infer welfare on sets that do not contain behavioural effects and are also different from the binary sets which are often considered as benchmark (Manzini et al., 2010), (Agranov & Ortoleva, 2017). Furthermore, since certain methods such as **MS**, **TC** and **EIG** are based on **CRP**, sets with 3 and 4 alternatives provide evidence on the capacity of these methods to break cycles in the **CRP** relation.

After the choice elicitation part subjects were asked, non-incentivized, to rank the four MAIN alternatives. No indifferences were permitted, hence the reported

[11]By Impatience we intend the violation of discounting models. The term "impatience" has been used by Fishburn & Rubinstein (1982) to denote Axiom A3.

[12]Given the high number of questions we apply "structural randomization". Namely, we divide questions into groups by similarity and then we completely randomize with the constraints that similar questions could not appear clustered together.

[13]For, example these comparisons are possible in Multiple Price Lists designs that are common in the literature of structural estimation of risk and time parameters (Andersson et al., 2016), (Andersen et al., 2008).

Table 2.1. **List of Main Delayed Payment Plans**

ALTERNATIVES	MONTHS				
	0	**3**	**6**	**9**	**12**
One Shot (OS)	160	0	0	0	0
Decreasing (D)	110	50	25	0	0
Constant (K)	50	50	50	50	0
Increasing (I)	0	15	40	170	0

Table 2.2. **List of Main Lotteries**

ALTERNATIVES	TOKEN		PROBABILITIES		EV
Degenerate (D)	50	0	1	0	50
Safe (S)	65	25	0.8	0.2	57
Fifty-Fifty (50)	90	25	0.5	0.5	57.5
Risky (R)	300	5	0.2	0.8	64

welfare relation is always a linear order.[14] Subsequently, subjects filled a questionnaire containing questions about the comprehension of the experimental design and criteria of choice in both Time and Risk. The questionnaire is presented and analysed in Appendix B.3. Finally, two well-known tests of cognitive abilities were presented: (i) Frederick Test - (Frederick, 2005); (ii) a selection of ten Raven matrices. Response times were collected for each question in the choice elicitation part and the cognitive abilities tests.[15]

The average reward was about 19 pounds per subject and the experiment lasted on average 1:15 hours. The reward was measured in Token with an exchange rate of 1:10 for lotteries and 1:20 for delayed payment plans. Subjects received no feedback about their earnings during the experiment. At the end of the experiment computers randomly picked from chosen delayed payment plans and lotteries, this latter was played out, and in the last screen informed subjects of their earnings in each part.

All sessions were conducted at the University of St. Andrews between June and September 2019. Subjects were recruited voluntarily among undergraduate and

[14]A linear order is a complete, transitive and antisymmetric binary relation.

[15]Since this experiment is part of a larger project, the analysis of cognitive abilities, response times and structural axioms is treated in a compendium paper (Caliari, 2020).

postgraduate students. Eleven sessions were run for a total of 145 subjects. No subject participated in more than one session. The earnings had been paid via bank account at the end of the experiment and in successive dates in the future as specified both by the instructions and by the experimenter. The experiment was completely anonymous and all subjects signed a consent form where they agreed in providing UK bank account number and sort code.

2.3 Results

2.3.1 CRP and BR

We begin showing the main result of the chapter. Table 2.3 presents the identification power of **CRP** and **BR** as the fraction of subjects for whom the methods can correctly identify either the reported best element or the entire welfare relation. As mentioned previously, both methods are founded on standard revealed preference, however the latter does not satisfy IR. **CRP** performs significantly better along all dimensions both in Time and Risk. Notably, **BR** is a lower bound for the identification since when a violation is observed data are simply ignored. This means that the difference is performed on subjects that violate WARP and therefore is not trivial.

Table 2.3. **CRP and BR - Identification**

	TIME			RISK		
METHODS	WRI	UI	EI	WRI	UI	EI
CRP	0.61	0.87	0.88	0.24	0.59	0.61
BR	0.42	0.59	0.74	0.06	0.14	0.43

NOTES -- **CRP** is the counting revealed preference method; **BR** denotes Bernheim & Rangel method. The numbers represent the fraction of subjects for whom the two welfare methods provide the following three identification: (1) "WRI" - Welfare Relation Identification and it refers to the unique identification of the entire reported welfare relation; (2) "UI" - Unique Identification of the reported best element; (3) "EI" - Expected Identification of the reported best element.

2.3.2 Premise: do individuals consistently reveal welfare?

Figure 2.2 presents the distribution of WARP violations in Time, Risk and random behaviour.[16] For each subject i, WARP violations are determined as the number of

[16]Two comments on random behaviour. First, given that the questions in Time and Risk were slightly different, random subjects may have different numbers of violations; however, the difference

cycles of length 2 in the graph of revealed preference.

$$WARP_i = \sum_{x,y} C_{xy} \cdot C_{yx}$$

Two observations catch the eye: (i) subjects violate WARP less in Time than in Risk and the difference in mean is statistically significant (t-test, $p = 0.000$); (ii) subjects do not behave randomly, again significantly (t-test, $p = 0.0000$).

The difference is not based only on the presence of a higher number of rational individuals in Time. If we restrict our test on those subjects that violate WARP at least once we find that the difference in mean is still highly significant (t-test, $p = 0.0002$). This suggests a fundamental difference in the behaviour of the agents in the two environments.

The suspicions are confirmed in Figure 2.3 where we show a scatter plot of the number of WARP violations. As the reader may notice the correlation is very low and driven mainly by a small fraction of consistent individuals. Given this preliminary evidence, we will treat Time and Risk separately in both consistency and preference elicitation analysis.

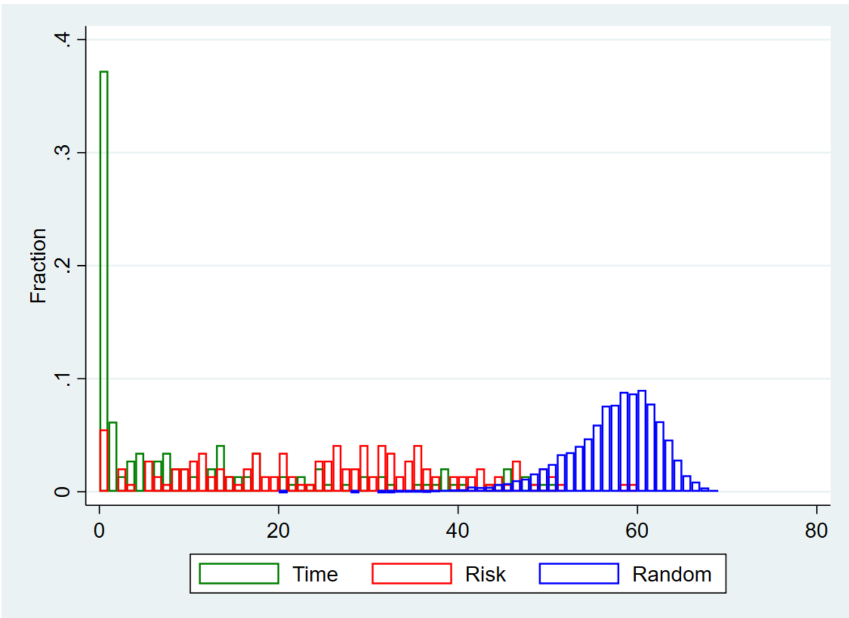


Figure 2.2. **Distribution of the violations of WARP.**

is negligible. Second, in order to provide a fair comparison we focus solely on the MAIN alternatives since they account for the vast majority of subjects' choices.

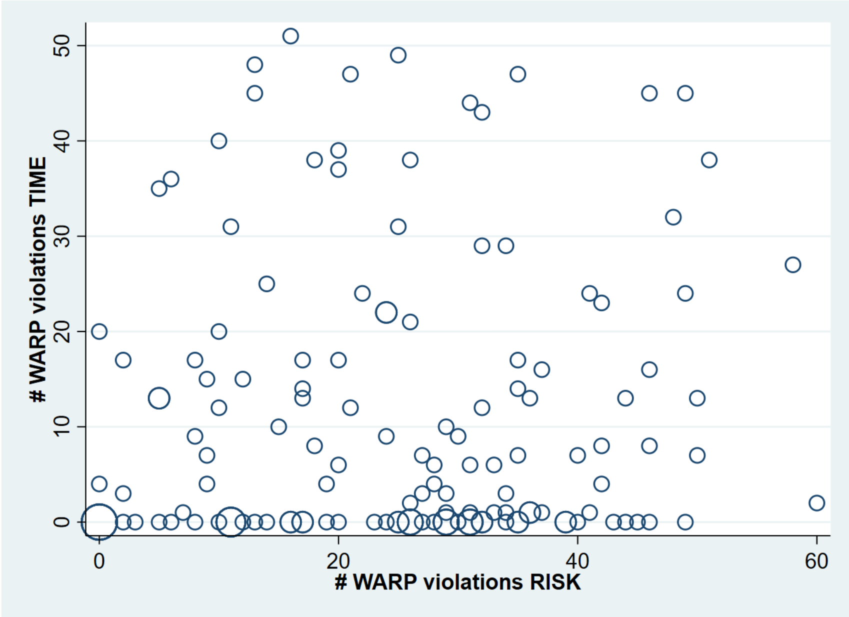


Figure 2.3. **Scatter plot of the violations of WARP.**

2.3.3 Consistency across the dataset

The first question that we investigate is: do different choice problems imply different levels of consistency? An answer will be crucial to draw a connection between consistency and welfare revelation. We divide the dataset into three parts called MAIN, AD, and BIG. The MAIN sets have been described in the previous section. AD refers to four sets doomed by asymmetric dominance, while BIG refers to the five (six in Risk) sets with more than eight elements.

Unfortunately, a simple comparison of the number of WARP violations across the dataset does not apply because this measure depends on the number and structure of the questions under scrutiny. In other words, we face the problem of: "... comparing the power of potentially different experimental designs. For a given choice setting, some experimental designs may be more likely to reveal violations of GARP than others." - Andreoni et al. (2013). The problem can be rephrased as follows: suppose one subject makes 10 inconsistent choices among 40 binary choices while another subject makes 10 inconsistent choices among 30 ternary choices. How can we compare these subjects in terms of consistency?

A standard approach in evaluating consistency of individuals given different experiments is to compare them with random behaviour - see Becker (1962) and

Bronars (1987). Some recent applications are Beatty & Crawford (2011) and Echenique et al. (2011). We address the problem constructing an index of consistency or "power index". We adopt the approach of perturbing a data generating process to create inconsistencies and compare the magnitude of the perturbation across domains.

As data generating process we build on the logit model as follows: let $A = \{x, y, z, w\}$ be the set of MAIN alternatives ordered by a linear order $\succ$ and u a utility function with $u(i) = u(j) + 1$ with $i, j \in A$ being consecutive elements in $\succ$. Note that, only differences in utility are important;[17] however, the parameter identification is not invariant to positive affine transformations of u (not cardinal). The standard logit formula is the following:

$$p(x, A) = \frac{e^{u(x)}}{\sum\limits_{y \in A} e^{u(y)}}$$

As in Train (2009)[18] we can modify the logit formula using a scale parameter λ connected to the variance of the unobserved error (a subject who chooses randomly behaves as if $\lambda = \infty$ but given our parameters for $\lambda \approx 5$ we substantially observe random behaviour); such that the formula becomes:

$$p(x, A) = \frac{e^{\frac{u(x)}{\lambda}}}{\sum\limits_{y \in A} e^{\frac{u(y)}{\lambda}}}$$

The parameter λ can be also interpreted as the cost of acquiring information regarding the utility of the elements, e.g. Caplin & Dean (2015) and Fudenberg et al. (2015).

We run a Monte Carlo simulation to estimate the parameter λ that match the average number of violations of WARP that the subjects make in the different part of the dataset. We only consider the MAIN alternatives since, as presented in Table 2.4

[17] Since in some part of the dataset the domain is not symmetric, namely some alternatives are more present than others. We adopt the convention of setting the utility difference of D and I (respectively S and R) equal to two. This is based on the fact the most of the subjects indicated in the ordinal ranking that these alternatives are divided by two positions; in particular, either $OS \succ D \succ K \succ I$ or $I \succ K \succ D \succ OS$. We also ignore confounding alternatives since they account for a marginal part of the choice distribution in any sets where MAIN alternatives are also present.

[18] An example of maximum likelihood estimate of the paraemter λ can be found in McKelvey & Palfrey (1995). They show that in a game theoretical experimental (quantal response equilibria) setting subjects tend, with experience, to make less noisy choices.

and 2.6, most of the violations, and choices, regard these alternatives.[19] Importantly this is not an estimation exercise (we do not believe that, when aggregated, subjects can be studied using a logit model). We provide an intuitive index that can be used for meaningful comparisons across domains. Given the strong assumptions made we also report the percentage of rational individuals and the standard deviation of our logit simulations such that the reader may have an idea of how close they are to the real data. We now present and comment on the consistency analysis in Time and Risk.

Time

The first part of Table 2.4 shows the mean and standard deviation of the distribution of WARP violations within different parts of the dataset, as well as the percentage of rational individuals, namely those with zero violations. In the second part, we present the logit index. The data show that AD questions present a relatively higher number of violations ($\lambda = 0.787$). The difference between BIG ($\lambda = 0.555$) and MAIN sets ($\lambda = 0.515$) is instead very small. To understand the importance of the λ measure, the reader may note that AD sets have both a low number of WARP violations and a high number of rational subjects.[20]

[19] This result is evidenced by the small difference between the violation in ALL** and ALL datasets. This assumption is conservative; in fact, in AD or BIG sets the identification of the parameter λ is lower than it would be.

[20] The simulation in AD and BIG sets ignores dominated alternatives. In the former we focus on four binary sets of the type $\{D, I\}$ with the assumption of $u(D) - u(I) = 2$, or vice versa. The assumption is based on the reported ranking of the high majority of individuals.

Table 2.4. **WARP Violations I - Time**

	BIG	AD	MAIN	ALL**	ALL
Mean	1.4897	0.8138	1.9586	10.0621	11.2621
Std	1.9189	1.4577	2.9009	14.2099	14.5263
Rational	54%	75%	59%	48%	37%
Logit - λ	0.555	0.787	0.515	0.569	-
Logit - std	1.6406	1.3738	2.0355	7.768	-
Logit - Rational	48%	74%	40%	16%	-

NOTES -- The mean of WARP violations is reported for different parts of the dataset: "BIG" denotes sets with more than 8 elements; "AD" denotes sets with potential asymmetric dominance effect; "MAIN" denotes the 11 non-empty subsets of the four main alternatives; "ALL" denotes the entire dataset. ALL** refers to WARP violations in the entire dataset that regard only the four main alternatives. We also report the following statistics: the information parameter of a logit model that match the data mean, the standard deviation and percentage of rational subjects in the resulting distribution.

Table 2.5. **WARP Violations II - Time**

	MAIN/BIG	MAIN/AD	BIG/AD
Mean	3.2897	1.3724	1.9172
Std	4.6682	2.5568	2.8052
Rational	56%	73%	56%
Logit - λ	0.515	1.062	1.124
Logit - std	2.8921	2.0696	2.508
Logit - Rational	34%	66%	59%

NOTES -- The mean of WARP violations is reported between difference domains: "MAIN/BIG" denotes violations observed between MAIN and BIG sets; "MAIN/AD" denotes violations between MAIN and AD sets; "BIG/AD" denotes violations between BIG and AD sets. These numbers are calculated, for instance, taking the total number of violations on MAIN and BIG sets and subtracting the violations within the two domains.

Two observations are worth noting. First, higher is the number of sets and worse is the logit approximation to the data. For instance, on the entire dataset, we should observe 16% of rational subjects while we observe 48% and the standard deviation is also significantly higher. Second, the coefficient of variation is everywhere above one. These observations suggest that there are, at least, two different groups of subjects: one rational and the other irrational. Importantly this latter has been shown to behave not randomly.

Table 2.5 reports the number of violations of WARP between different domains. For instance, when **x** is chosen over **y** in a MAIN set and **y** over **x** in a BIG set. The results show that not only the level of rationality is similar within MAIN and BIG sets (Table 2.4) but also the types of violations are similar ($\lambda = 0.515$). On the other hand, AD sets present a different behaviour from both MAIN and BIG sets (resp. $\lambda = 1.062$ and $\lambda = 1.124$). Notice that to match the number of WARP violations between AD sets and the other domains we would require a level of perturbation higher than all levels within the domains. Furthermore, Table 2.5 confirms the presence of at least two groups of individuals since the standard deviation of the logit simulations is everywhere below the standard deviation in the data.

Risk

Table 2.6 reports the results regarding WARP violations within domains in Risk. The number of violations is, on average, higher than in Time across all the domains and everywhere significantly (t-test, $p = 0.0000$ in MAIN, $p = 0.015$ in AD, $p = 0.0000$ in BIG). In this case, the comparison between Time and Risk is meaningful given the approximate symmetry of the datasets. This evidence suggests that the difference in behaviour between the two environments is not due to a particular incidence of behavioural effects. The difference in the shape of the distribution, expressed in Figure 2.2, is confirmed by the coefficients of variation. If in Time they were everywhere above one, confirming that left skewness is a common property across domains, in Risk they are almost everywhere below one, confirming the generality of the uniform shape of the distribution. Surprisingly, Table 2.6 shows that in BIG sets ($\lambda = 0.756$) subjects appear more rational compared to both MAIN ($\lambda = 1.009$) and AD sets ($\lambda = 1.003$).[21][22] Data also confirm that when the number of sets increases the percentage of rational subjects becomes higher than the one in the logit simulation.

[21] In a compendium paper (Caliari, 2020), we report evidence of deliberate randomization as modelled by Cerreia-Vioglio et al. (2019) and reported by Agranov & Ortoleva (2017). Behavioural effects may reduce the capacity of subjects to deliberately randomize, therefore reducing WARP violations.

[22] This evidence may be related with attention models such Masatlioglu et al. (2012), Manzini & Mariotti (2014), Lleras et al. (2017) and Cattaneo et al. (2018), and could confirm previous experiments such as Iyengar & Kamenica (2010). On the contrary, models that assume more uniform stochastic choice in BIG sets such as Fudenberg et al. (2015) and Frick (2016) seem to be not backed by the data.

Table 2.6. **WARP Violations I - Risk**

	BIG	AD	MAIN	ALL**	ALL
Mean	4.6690	1.2621	4.9862	21.7172	24.6552
Std	2.8700	1.4955	3.4237	13.6314	14.3170
Rational	15%	54%	14%	8%	6%
Logit - λ	0.756	1.003	1.009	0.774	-
Logit - std	2.9899	1.5672	2.6606	9.6465	-
Logit - Rational	22%	60%	7%	2%	-

NOTES -- The mean of WARP violations is reported for different parts of the dataset: "BIG" denotes sets with more than 8 elements; "AD" denotes sets with potential asymmetric dominance effect; "MAIN" denotes the 11 non-emtpy subsets of the four main alternatives; "ALL" denotes the entire dataset. ALL** refers to WARP violations in the entire dataset that regard only the four main alternatives. We also report the following statistics: the information parameter of a logit model that match the data mean, the standard deviation and percentage of rational subjects in the resulting distribution.

Table 2.7. **WARP Violations II - Risk**

	MAIN/BIG	MAIN/AD	BIG/AD
Mean	9.0276	2.3241	2.3862
Std	5.9224	2.7267	2.5888
Rational	14%	44%	40%
Logit - λ	0.688	1.125	1.581
Logit - std	4.7995	2.2564	2.3141
Logit - Rational	9%	37%	32%

NOTES -- The mean of WARP violations is reported between difference domains: "MAIN/BIG" denotes violations observed between MAIN and BIG sets; "MAIN/AD" denotes violations between MAIN and AD sets; "BIG/AD" denotes violations between BIG and AD sets. These numbers are calculated, for instance, taking the total number of violations on MAIN and BIG sets and subtracting the violations within the two domains.

Table 2.7 reports a higher similarity in the behaviour of subjects in MAIN and BIG sets ($\lambda = 0.688$) compared to both MAIN/AD and BIG/AD sets (resp. $\lambda = 1.125$ and $\lambda = 1.581$). It is particularly interesting to notice the extremely high logit index associated with violations between BIG and AD sets. Speculations would lead us to conjecture that choice overload and asymmetric dominance, although both in the family of behavioural effects, have very different implications on the consistency of behaviour in choice among lotteries.

2.3.4 Identification of reported welfare

This subsection contains the main results of Chapter II. We measure the power of identification of different welfare methods in both Time and Risk using ALL dataset, MAIN sets, and BINARY sets. This latter is considered as a benchmark to understand how much information can be extracted from a dataset that does not present any potential behavioural effect. Two results emerge in both Time and Risk: (1) methods that satisfy IR performs significantly better than **BR**; (2) the identification power of methods that satisfy IR improves when more data are collected. This result, as expected, is reversed in **BR**.

Our identification exercise is threefold. Firstly, we uniquely identify the reported best element. Secondly, since **BR** is a conservative approach, it is reasonable to imagine that this method performs better in a set identification exercise; namely when the reported best element is in the set of maximal elements. We assume that a risk-neutral policy maker has to pick from the set of maximal elements endowed with a uniform distribution. Given this assumption, we perform an expected identification exercise. Finally, we uniquely identify the entire reported welfare relation.

Let N be the set of subjects and $f_i(D)$ be the preference elicited by the welfare method f given the choices of subject i over the dataset D. The reported welfare relation by subject i is denoted as $\mathsf{REP}_i(\succ)$. The proportion of correctly identified subjects given the three approaches is as follows:

- Unique Identification [**UI**]:

$$\frac{\#\{i \in N : \max[\mathsf{REP}_i(\succ)] = \max[f_i(D)]\}}{\#N}$$

- Expected Identification [**EI**]:

$$\frac{\sum_{i \in N : \max[\mathsf{REP}_i(\succ)] \in \max[f_i(D)]} \frac{1}{\#\{\max[f_i(D)]\}}}{\#N}$$

- Welfare Relation Identification [**WRI**]:

$$\frac{\#\{i \in N : \mathsf{REP}_i(\succ) = f_i(D)\}}{\#N}$$

Note that, the reported welfare relation is necessarily asymmetric. Hence, methods that map into linear orders such as **SEQ** or **EIG** are theoretically favoured in the identification of the entire welfare relation. To solve this issue we also investigate how close methods are to identify reported welfare relation even when these are not perfectly identified. The similarity of solutions is measured using the sum over all subjects of:

- Symmetric Difference [**SD**] between the resulting binary relations and the reported order. The symmetric difference $\triangle$ between two binary relations R_1, R_2 is defined as follows[23]:

$$R_1 \triangle R_2 = (R_1 \setminus R_2) \cup (R_2 \setminus R_1)$$

- Reverse Asymmetry [**RA**], denoted here as $\bigtriangledown$, is defined as the number of times the asymmetric part of the reported order is reversed. Namely, given two asymmetric binary relations P_1 and P_2:

$$P_1 \bigtriangledown P_2 = |\{(x,y) \in P_1 : (y,x) \in P_2\}|$$

Using both measures is crucial. The symmetric difference considers equally the symmetric and asymmetric part of the binary relation, hence punishing coarse methods such as **BR**. The "reverse asymmetry" measure allows us to disentangle those differences that are in principle worse; namely when a subject reports **x** better than **y** but the method ranks **y** better than **x**. This measure punishes particularly methods that map in linear orders such as **EIG** and **SEQ**; while the conservative nature of **BR** creates a lowest bound. This analysis, together with the three identification exercises, provides a comprehensive picture of the identification power of each method.

[23] For instance, let $R_1 = \{(x,y),(y,x),(y,z),(x,z)\}$ and $R_2 = \{(x,y),(y,z),(z,y),(x,z)\}$ we have $R_1 \triangle R_2 = |\{(y,x),(z,y)\}| = 2$.

The empirical role of Robustness and Stability

Before introducing the identification results using experimental data, it is important to highlight how Robustness [ROB] and Stability [ST] can affect the results. To understand the issue the reader can refer to the connection that has been drawn between ROB, ST and continuity (see Appendix A). In particular, we would like to measure how a method is sensitive to perturbations in the choice space. In order to do that, we use the above introduced concept of Reverse Asymmetry. We analyse the set of single-valued choice functions (see Chapter 1 for a formal definition) with four elements on our MAIN sets. The dimension of this space is of 20736 choice functions. Then, we adopt a Leave-one-out test for continuity. We eliminate one set from the dataset and observe how the inference of welfare relations changes. Below we report the average number of RA per choice function for each method.

	BIN	TER	QUAT
CRP	0	0	0
MS	0.0187	0.1076	0.2384
TC	0	0.0584	0.1493
EIG	0.2521	0.6357	1.0729
CC	0	0	0
SEQ	×	×	×
BR	0	0	0

Figure 2.4. **Degree of Continuity of methods**

From the figure above shows, in line with the theory, that **CRP**, **CC** and **BR** are stable methods. The remaining methods instead are ranked **TC**, **MS**, **EIG** with the latter being the most discontinuous. As it will be clear in the next section, higher is the level of discontinuity and higher is the probability that new information will reverse the judgement. Hence, on one hand, in contexts where few inconsistencies are observed and the behaviour of the subjects is homogeneous across domains (Time preferences) a discontinuous method may wrongly overestimate rare mistakes. On the other hand, when more inconsistencies are observed and the behaviour of subjects is less homogeneous a discontinuous method may still be able to provide a point estimate and to be sensitive to new information (Risk preferences).

Time

Table 2.8 shows that methods that satisfy IR perform significantly better than **BR** both uniquely ($\approx 30\%$) and in expectation ($\approx 15\%$). It is crucial to notice that **BR** is a lowest bound in the identification exercise since it identifies only those subjects that rationally reveal their best element. Therefore, the 30% gap is not trivial because it is performed on irrational individuals.

Table 2.8. **Unique and Expected Identification - Time**

		UNIQUE			EXPECTED		
METHODS		**ALL**	**MAIN**	**BINARY**	**ALL**	**MAIN**	**BINARY**
IR & CRP	**CRP**	0.87	0.81	0.77	0.88	0.84	0.77
	MS	0.87	0.81	0.79	0.88	0.85	0.80
	EIG	0.87	0.83	0.81	0.87	0.83	0.81
	TC	0.88	0.81	0.77	0.88	0.83	0.77
IR	**CC**	0.81	0.83	0.77	0.84	0.86	0.81
No-IR	**SEQ**	-	0.83	-	-	0.83	-
	BR	0.59	0.67	0.77	0.74	0.79	0.77
	OW	0.89	-	-	0.89	-	-

NOTES -- On the left we show the portion of subjects for whom each method uniquely identify the reported best element. On the right, the expected portion of subjects for whom each method identify the reported best element. The measure is expected because for some subjects methods may set identify the best element; in these cases we assume to pick uniformly from the set of identified elements.

The power of identification for methods that satisfy IR is increasing in the number of sets in the dataset which suggests that individuals reveal information about welfare along all the dataset. Only exception is **CC**. We interpret this as evidence in favour of the importance of standard revealed preference as a foundation for welfare methods.

Finally, **SEQ** performs particularly well; the difference is only 4-6%. The reason is that the best element of **SEQ** is the one chosen from the set with all the four MAIN alternatives. It turns out this choice is a good predictor of the reported best element, although the two elicitations are not equivalent.

Table 2.9. **Iden. Welfare Relation, SD and RA - Time**

METHODS		WRI ALL	WRI MAIN	WRI BINARY	SD & RA ALL SD	ALL RA	MAIN SD	MAIN RA	BINARY SD	BINARY RA
		-	-	-	SD	RA	SD	RA	SD	RA
IR & CRP	**CRP**	0.61	0.57	0.59	180	78	191	73	220	110
	MS	0.62	0.59	0.61	182	82	188	76	218	88
	EIG	0.54	0.60	0.61	222	111	208	104	218	109
	TC	0.61	0.58	0.59	180	73	188	68	234	71
IR	**CC**	0.54	0.58	0.59	214	91	186	74	218	78
No-IR	**SEQ**	-	0.60	-	-	-	194	97	-	-
	BR	0.42	0.50	0.59	264	45	226	54	220	110
	OW	0.66	-	-	170	85	-	-	-	-

NOTES -- On the left we show the portion of subjects for whom each method uniquely identify the entire reported welfare relation. On the right, "SD" and "RA" denote respectively symmetric difference and reverse asymmetry.

Table 2.9 reports the identification of the entire welfare relation. We present it together with symmetric difference and reverse asymmetry measures. We confirm that methods that satisfy IR perform better than **BR** by 10-15% when we look at the left part of Table 2.9, namely the percentage of total subjects that have been uniquely identified. However, as mentioned before, this observation is not enough to judge the methods. For instance, the performances of **SEQ** and **EIG** are positively biased by the feature that they map into linear orders. Since the reported preference relations are linear orders by construction, the probability that this latter are uniquely identified is higher. As mentioned in the previous subsection, we use SD and RA to measure the distance between the reported and elicited binary relations.

First, we observe that as theoretically predicted **BR** provides on one hand a lower bound on RA given its cautious approach described by its stability and infinitely non-robustness. On the other hand, it provides an upper bound on SD given its coarseness. Considering the other methods in comparison with **BR**(SD,RA) = $(264, 45)$, we can see that **SEQ** and **EIG** are significantly outperformed both in SD and in RA by both **MS** and **TC**.

The monotonicity of the identification power in the size of the dataset is not straightforward here. However, if we observe the SD of methods that satisfy IR we notice that it is decreasing for any method apart from **EIG** and **CC**. The latter result confirms the weakness of IR and the necessity of focusing on methods based on **CRP**. The former instead is based on the excessive non-robustness of **EIG**, and in particular, the excessive weight that this method poses on observations from BIG sets, that by construction, induce a higher change in **CRP** as shown in Figure 2.4. Furthermore, as we will show in Section 2.3.7, AD sets in Time preferences are not important to infer the welfare relations, hence when added they may cause a problem in the inference of non-robust methods. Finally, Binary sets are shown to be very important and therefore a highly non-robust method such as **EIG**, may lose power of identification when less important observations are added.

Risk

Table 2.10 reports the identification results in Risk. Data show that methods that satisfy IR perform significantly better than **BR** both uniquely (50%) and in expectation (20%). We also confirm that the power of identification is generally (note that **CC** is still an exception) increasing in the size of the dataset.

The choice from the set of MAIN alternatives is again a good predictor of the reported best element since the loss of **SEQ** is only 4-8%.

Table 2.10. **Unique and Expected Identification - Risk**

		UNIQUE			EXPECTED		
	METHODS	**ALL**	**MAIN**	**BINARY**	**ALL**	**MAIN**	**BINARY**
IR & CRP	**CRP**	0.59	0.52	0.42	0.61	0.59	0.42
	MS	0.59	0.52	0.46	0.61	0.60	0.50
	EIG	0.61	0.61	0.51	0.61	0.61	0.51
	TC	0.61	0.51	0.42	0.62	0.55	0.42
IR	**CC**	0.55	0.56	0.42	0.57	0.61	0.50
No-IR	**SEQ**	-	0.55	-	-	0.55	-
	BR	0.14	0.25	0.42	0.43	0.49	0.42
	OW	0.63	-	-	0.63	-	-

NOTES -- On the left we show the portion of subjects for whom each method uniquely identify the reported best element. On the right, the expected portion of subjects for whom each method identify the reported best element. The measure is expected because for some subjects methods may set identify the best element; in these cases we assume to pick uniformly from the set of identified elements.

The left part of Table 2.11 again shows that methods that satisfy IR outperform **BR** in the entire identification exercise (15-20%). We also confirm that **SEQ** and **EIG** performances are only apparently good; in fact, when controlled for RA measure, and normalizing for the RA measure of **BR**, we see that they perform worse than **MS** by respectively 25% and 17%. Looking at the SD we observe that **SEQ** is clearly outperformed while **EIG** performance is relatively superior to the one in Time preferences. We also note that the monotonicity of the identification in the size of the dataset is confirmed everywhere IR is satisfied, hence also in **EIG**. This result suggests that AD and BIG sets add very valuable information in this case, and on the other hand that Binary sets are not as important as they are in Time preference to predict the reported preference relation. Both these results are confirmed in Section 2.3.7.

Table 2.11. **Iden. Welfare Relation, SD and RA - Risk**

		ENTIRE IDEN.			SD & RA					
		ALL	**MAIN**	**BINARY**	**ALL**		**MAIN**		**BINARY**	
METHODS		-	-	-	SD	RA	SD	RA	SD	RA
IR & CRP	**CRP**	0.24	0.19	0.20	436	186	455	168	556	278
	MS	0.24	0.20	0.21	440	190	452	179	569	241
	EIG	0.30	0.27	0.23	446	223	448	224	576	288
	TC	0.24	0.19	0.20	434	182	446	157	570	184
IR	**CC**	0.21	0.19	0.20	453	200	452	185	569	218
No-IR	**SEQ**	-	0.25	-	-	-	478	239	-	-
	BR	0.06	0.10	0.20	592	86	545	115	556	278
	OW	0.32	-	-	421	210	-	-	-	-

NOTES -- On the left we show the portion of subjects for whom each method uniquely identify the entire reported welfare relation. On the right, "SD" and "RA" denote respectively symmetric difference and reverse asymmetry.

2.3.5 Optimal Weighting

So far, we have shown that methods that satisfy Informational Responsiveness and are based on revealed preference guarantee better performances in the identification exercises. From now on, we aim to solve two possible drawbacks. First, the reader knows only a relative measure of the performance of these methods. Namely, that they perform better than the alternative ones. However, we aim to provide a more general measure of performance confronting them with a data-driven welfare method the we call Optimal Weighting method [**OW**]. Second, so far we have tested

IR only indirectly, namely using methods that do or do not satisfy it. **OW** allows us to test if IR directly binds for the identification exercises. In words, does a method that differently weight observations in order to optimize the identification exercise, assigns strictly positive weights to all observations where **x** is chosen and **y** is available, therefore satisfying IR?

To define **OW** we divide the dataset in five parts: binary sets [B], ternary sets [T], quaternary set [Q], sets with asymmetric dominance [AD], big sets [BIG]. For each part the revealed preference is collected creating, for each $x, y \in X$, a vector $\mathbf{C}_{xy} = (C^B_{xy}, C^T_{xy}, C^Q_{xy}, C^{AD}_{xy}, C^{BIG}_{xy})$. The weights vector is $\mathbf{w} = (w_B, w_T, w_Q, w_{AD}, w_{BIG})$. We define the method OW as follows:

$$xR^D_{\mathbf{OW}}y \text{ if and only if } OW_{xy} \geq OW_{yx}$$

where $OW_{xy} = \sum\limits_{l \in \Gamma} w_l C^l_{xy}$ and $\Gamma = \{B, T, Q, AD, BIG\}$.

Weights are calculated optimizing the sum of two measures: (1) expected identification of maximal element [**EI**]; (2) unique identification of the entire welfare relation [**WRI**]. To recall, the former measures the expected number of subjects for whom the method can identify the reported best element; the latter measures the number of subjects for whom the method uniquely identify the entire reported welfare relation.[24] The optimization problem is as follows:

$$\max_{\mathbf{w} \in [-0.4, 1]^5} \mathbf{EI} + \mathbf{WRI}$$

where for each subject i:

$$xR^D_{f_i}y \quad \Leftrightarrow \quad \mathbf{w} \cdot \mathbf{C}_{xy_i} \geq \mathbf{w} \cdot \mathbf{C}_{yx_i}$$

Two main features of the **OW** method allows us to understand its relevance. First, the objective function is a distance between the reported and the elicited preference relations. Therefore, it would be as if we knew the reported relations (data-driven) and we are trying to get closer to them optimizing on the importance of the choices made by the individual in different parts of the dataset. Given the generality of the weighted average adopted, **OW** will have a better performance compared to

[24]The optimality problem is performed using different objective functions in Section 2.3.7

the other methods that are instead not data-driven. Second, the weights attached to different parts of the dataset may be negative. Consider the following example, from the MAIN sets, a subject always chooses x when available, y if x is not available, and in the binary set $\{z, w\}$ he chooses z. Then, he reports $x \succ y \succ w \succ z$. In this case, it may be that since x, y are clearly best, binary sets receive a small negative weight that guarantees $w \succ z$, and does not change the other preferences.

2.3.6 Completeness of the methods

In this section, we compare the identification results of welfare methods with the data-driven method and refer to the distance between them as the completeness of the methods. We borrow the term "completeness" from Fudenberg et al. (2019). In their paper, the authors use machine learning to measure the amount of variation in the data that a theory can capture. Their notion of completeness aims to answer the following question: "How close is the performance of a given theory to the best performance that is achievable in the domain?" (Fudenberg et al., 2019). In our framework, we define completeness, denoted as $\texttt{Com}(f)$ for some welfare method f, as:

$$\texttt{Com}(f) = \frac{\varepsilon(f_L) - \varepsilon(f)}{\varepsilon(f_L) - \varepsilon(f_U)}$$

where $\varepsilon(f_L)$ is the proportion of non-identified subjects by the method that defines a lower bound on the domain; $\varepsilon(f_U)$ is the best achievable residual proportion and $\varepsilon(f)$ is the residual proportion of the model under study. In our framework, we set $f_L = \textbf{BR}$ and $f_U = \textbf{OW}$. Table 2.12 shows the completeness of the methods using ALL sets across different types of identification procedures in both Time and Risk.

Table 2.12. **Completeness of the methods**

	TIME			RISK		
METHODS	UI	EI	WRI	UI	EI	WRI
CRP	0.93	0.93	0.79	0.92	0.90	0.69
MS	0.93	0.93	0.83	0.92	0.90	0.69
EIG	0.93	0.86	0.50	0.96	0.90	0.92
TC	0.95	0.93	0.79	0.96	0.95	0.69
CC	0.74	0.66	0.50	0.84	0.70	0.58
SEQ	0.81	0.59	0.75	0.84	0.60	0.73
BR	0.00	0.00	0.00	0.00	0.00	0.00
OW	1.00	1.00	1.00	1.00	1.00	1.00

NOTES -- This table reports the completeness of all methods in cases of unique (UI), expected (EI) and entire (WRI) identification procedures.

Since **BR** and **OW** are respectively lower and upper bound for our identification analysis they take respectively value zero and one. Methods that satisfy IR and are based on the revealed preference approach have generally higher completeness than other methods. Note that, even though we do not report completeness for the measures of symmetric difference and reverse asymmetry in the entire identification approach, that favours **SEQ** over other methods, there always exists at least a method among those that satisfy IR and are based on revealed preference that is more complete than **SEQ**.

2.3.7 A direct test of Informational Responsiveness

We propose a direct test for IR that exploits the data-driven method **OW**. We focus on the family of methods that are weighted sums of **CRP**, depending perhaps on the sets in which the choice has been observed. If each choice receives a strictly positive weight independently from the set where the choice happened then IR is satisfied.[25] In this sense, our construction of **OW** allows us to test whether IR binds in an optimal identification problem.

We generalize our previous analysis where the convention was to optimize the sum of expected identification of the reported best element and unique identification of the entire welfare relation. In this section, we report results based on

[25]This implication is immediate. See Meyer & Mongin (1995) for a comprehensive study of affine aggregation.

six different objective functions. Before introducing them, some clarifications are needed. First, the optimization problem described in Section 2.3.5 may clearly have non unique results. Namely, there may be multiple set of weights that optimize the objective function. In such cases, we report the minimum and maximum weights for each part of the dataset such that there exists a system of weights that solve the optimization problem. Importantly, this does not imply that any vector of weights that is in the Cartesian Product of the intervals guarantees optimal identification. Second, as it may be clear from the example about negative weights at Section 2.3.5, if choices from a particular set are irrelevant then this set may receive positive, negative or a zero weight without changing the result. This observation has important consequences in the interpretation of the results. For instance, in Table 2.13 and Table 2.14, AD sets receive both negative and positive weights ($[-0.2, 1]$ in Time and $[-0.2, 0.9]$ in Risk) when we try to optimize on the identification of the best element (**UI** and **EI**). This is due to the fact that only two alternatives, in Time and Risk, are represented in AD sets and often these alternatives are not reported and chosen as best alternatives. Therefore, from this interval we cannot conclude anything about the importance of AD sets in eliciting preferences and we need to focus on the weights assigned to AD sets under the remaining four objective functions.

Time

Table 2.13 shows the intervals of weights that guarantee optimality for the six different objective functions. For completeness of information, we split the MAIN sets into three parts: Binary sets, Ternary sets, and Quaternary set.

Table 2.13. **Optimal Weights - Time**

	TIME				
IDENTIFICATIONS	**BIN**	**TER**	**QUA**	**BIG**	**AD**
UI	[0.6,1]	[0.2,1]	[0.1,0.2]	[0.6,1]	[-0.2,1]
EI	[0.6,1]	[0.2,1]	[0.1,0.2]	[0.6,1]	[-0.2,1]
WRI	[0.2,0.9]	[0.3,1]	[0.3,1]	[0.4,1]	[-0.2,-0.1]
SD	[0.5,0.8]	[0.6,1]	[0.4,0.8]	[0.4,0.7]	-0.2
SD & RA	0.6	0.6	0.6	0.6	-0.2
EI & WRI	0.9	1	0.4	0.8	-0.2

NOTES -- The table contains intervals of weights that optimize the identification of different objectives. "UI" and "EI" denote respectively unique and expected identification of the best element; "WRI" denotes entire welfare relation identification; "SD" and "RA" denote respectively minimization of the sum of symmetric difference and [two times] reverse asymmetry against the reported welfare relation; "EI & WRI" denotes the sum of EI and WRI. This latter is the one used along the paper to define OW.

We observe that strictly positive weights are associated to any part of the dataset apart from AD sets. This latter is found to be irrelevant in the identification of the reported best element (weights can be negative, zero, or positive), while they have negative weights when we identify the entire welfare relation. As above mentioned, the first result is expected, while the second result is somewhat surprising since it shows that subjects wrongly reveal their welfare in this part of the dataset. Nonetheless, it confirms the findings of Section 2.3.3, where we show that subjects are not only more irrational in these sets (Table 2.4); but also they have a different behaviour (Table 2.5) if compared to MAIN and BIG sets.

We also find that binary sets are particularly important throughout all the possible objective functions. This explains both the relatively good performance of methods on these sets (Table 2.8) and the fact that the identification power of **EIG** decreases in the size of the sets as observed in Table 2.9. This is due to the high weight put to bigger sets by the **EIG** method.

Risk

Table 2.14 shows that, in Risk, IR binds everywhere since strictly positive weights are attached to any domain. There are two exceptions. Firstly, AD sets are irrelevant when we focus only on the reported best element, but again this observation is not relevant from an empirical perspective. However, in the remaining cases, AD sets receive strictly positive weights, differently from Time, showing that in Risk behavioural effects such as attraction effect and compromise effect seem not to undermine the elicitation of preferences.

Table 2.14. **Optimal Weights - Risk**

	RISK				
IDENTIFICATIONS	**BIN**	**TER**	**QUA**	**BIG**	**AD**
UI	[-0.2,0]	[0.4,0.7]	[0.7,1]	[0.5,0.9]	[-0.2,0.9]
EI	-0.1	[0.3,0.8]	[0.5,1]	[0.4,0.9]	[-0.2,0.9]
WRI	[0.2,0.7]	[0.4,1]	[0.8,1]	[0.3,0.8]	[0.3,1]
SD	0.4	0.4	1	0.3	0.4
SD & RA	0.5	[0.4,0.5]	[0.8,1]	[0.4,0.5]	[0.4,0.5]
EI & WRI	[0.1,0.3]	[0.4,0.5]	[0.8,1]	[0.4,0.6]	[0.3,0.6]

NOTES -- The table contains intervals of weights that optimize the identification of different objectives. "UI" and "EI" denote respectively unique and expected identification of the best element; "WRI" denotes entire welfare relation identification; "SD" and "RA" denote respectively minimization of the sum of symmetric difference and [two times] reverse asymmetry against the reported welfare relation; "EI & WRI" denotes the sum of EI and WRI. This latter is the one used along the paper to define OW.

Secondly, when we focus only on the identification of the reported best element we observe that binary sets receive weakly negative weights. These weights are also strictly positive but close to zero in the other cases. This confirms the findings of previous sections. In fact, in Table 2.10 we find that methods perform poorly on binary sets. We also found (Table 2.11) that the **EIG** method has an increasing identification power in the size of the sets. And again in Table 2.11 we find that, throughout all methods, the differential of both symmetric difference and reverse asymmetry between binary sets, MAIN and ALL datasets is positive and significant.

The low importance of binary sets is striking. Especially, if we compare the weights associated with BIG sets where supposedly we should observe choice overload effect. This seems to suggest that, in Risk, the irrational behaviour in MAIN sets is mostly driven by binary sets.[26]

2.4 Conclusion

Using a novel experimental design, we test the hypothesis that Informational Responsiveness and Revealed Preference are necessary conditions for behavioural wel-

[26]This evidence suggests further research on attention in choice among gambles and it is in line with stochastic models such as Manzini & Mariotti (2014) and Cattaneo et al. (2018).

fare analysis. Firstly, we show that individuals repeatedly violate the Weak Axiom of Revealed Preference both in time and risk preferences. We develop a new index of rationality and show that inconsistency is a general phenomenon, namely it is common to sets with different cardinality and with or without behavioural effects. Secondly, we find that welfare methods that satisfy Informational Responsiveness and are based on a Revealed Preference approach perform significantly better in identifying both the best reported element and the entire reported welfare relation. The results are strong in both time and risk preferences and in any part of the dataset. We show that these welfare methods are more complete theories in the sense of Fudenberg et al. (2019). Finally, using an optimal weighting algorithm we directly test Informational Responsiveness. We show that subjects reveal welfare in all parts of the dataset. Therefore, we argue that welfare analysis should not ignore data doomed by behavioural effects but only eventually give different weights to such observations. Our analysis does not solve the elicitation problem entirely. The researcher's problem, though much simplified, still requires a correct evaluation of these welfare weights.

Chapter 3

Imperfect Discrimination and Stochastic Transitivity

3.1 Introduction

Imperfect discrimination is a widely studied feature of human cognition. In choice theory, it forms the behavioural foundations of deterministic models such as Luce's semiorder (Luce, 1956)[1] and stochastic models such as the Fechnerian model (Thurstone, 1927).[2]

In this Chapter, we analyse the behavioural consequences of imperfect discrimination when it is allowed to vary with the alternatives under scrutiny. As an illustrative example, we refer to the so-called "similarity hypothesis" (Tversky & Russo, 1969). The authors conjecture that more similar alternatives are easier to compare, and therefore, keeping fixed the utility of the alternatives, higher the similarity and easier is the choice of the best alternative. Tversky (1972) describes this idea in the famous Paris-Rome example. He imagines a decision maker than has to choose between a trip to Paris [**P**], a trip to Rome [**R**] and a trip to Paris with $1 bonus [**P+**]. He notices that a decision maker may find hard to decide between **P** or **P+** and **R**, but he would not have any doubts when deciding between **P+** and **P**. Choice situ-

[1] Luce (1956): "The nontransitiveness of indifference must be recognized and explained on any theory of choice, and the only explanation that seems to work is based on the imperfect powers of discrimination of the human mind whereby inequalities become recognizable only when of sufficient magnitude."

[2] McFadden (1980) wrote, referring to Thurstone (1927): "To accommodate the demonstrated inability of individuals to discriminate perfectly... utility is a random function."

ations of this kind induce particular properties of the resulting choices in both the deterministic and stochastic environment.

More specifically, we show that different notions of transitivity of preferences can be characterized using the idea of imperfect discrimination. In the previous example, assume a decision maker prefers **P+** to **P** but he is uncertain of his preferences (e.g. incomplete preference) between **P** and **R**. Assume also that there exists a measure of discrimination that can be interpreted as the uncertainty regarding the utility of the alternatives. One result from the deterministic choice literature states that if the uncertainty between **P+**, **R** is smaller than the sum of the uncertainties between **P+**, **P**, and **P**, **R** then the preference relation is transitive - (Nakamura, 2002). Assuming no imperfect discrimination between **P+** and **P**, it becomes easy to explain a situation in which the only clear preference is between **P+** and **P** (e.g. the uncertainty regarding the difference in utility between **P+**, **P** and **R** is high). If, for example, the decision maker prefers **P** to **R**, then he will also prefer **P+** to **R** due to the transitivity property.

Our main result is to generalize the above example to the stochastic choice literature. We provide a complete characterization of all the main notions of stochastic transitivity: Weak [WST], Moderate [MST] and Strong [SST]. These conditions are pervasive in the literature of stochastic choice models. The table below shows the connection between a series of models and each stochastic transitivity property.

Models that may violate WST	
Random Utility models	Marschak & Block (1960)
Attribute Rule	Gul et al. (2014)
Random Consideration Set Rule	Manzini & Mariotti (2014)
Random Consideration Choice Set Rule	Brady & Rehbeck (2016)
Dual Random Utility model	Manzini & Mariotti (2018)
Deliberate Randomization	Cerreia-Vioglio et al. (2019)
Focus, then compare	Ravid & Stevenson (2019)
Models that satisfy WST	
Item Invariant Additive Perturbed Utility	Fudenberg et al. (2014)
Gradual Pairwise Comparison Rule	Dutta (2020)
Models that satisfy MST	
Tversky EBA	Tversky (1972)
Menu Invariant Additive Perturbed Utility	Fudenberg et al. (2014)
Single-Crossing Random Utility Model	Apesteguia et al. (2017)
Moderate Utility	He & Natenzon (2018)
Bayesian Probit	Natenzon (2019)
Models that satisfy SST	
Fechnerian Model	Thurstone (1927) - Debreu (1958)
Luce model	Luce (1959)
Simple scalable model	Tversky & Russo (1969)
Additive Perturbed Utility	Fudenberg et al. (2015)
Symmetric Random Utility Model	Marschak & Block (1960)

We model imperfect discrimination as a property of pairs of alternatives in both a deterministic and stochastic model. We refer to the former as Binary Threshold model and to the latter as Binary Additive Perturbed Utility model. According to the former, the chosen elements from a binary set of alternatives $\{x,y\} \subseteq A$ are determined as follows:

$$c(\{x,y\}) = \{x \in \{x,y\} : u(y) - u(x) \leq \varepsilon(x,y)\}$$

where u is a utility function and $\varepsilon : A \times A \to \Re^+$ is a threshold function. In the stochastic model the probability distributions over binary sets are determined as follows:

$$p(\{x,y\}) = \underset{p \in \Delta(\{x,y\})}{\operatorname{argmax}} \sum_{z \in \{x,y\}} u(z)p(z) - \eta(x,y) \cdot c(p(z))$$

where again u is a utility function, c is a strictly convex function, and $\eta : A \times A \to \Re^{++}$ represents imperfect discrimination between pairs of alternatives.[3] Both ε and η are symmetric.

Some preliminary comments. Intuitively, the BAPU model can be seen as a refinement of the BT model in the sense that a stochastic choice function resulting from BAPU contains more information about both utility and imperfect discrimination than the choice correspondence resulting from BT. We will show this connection relying on the properties that characterize the two models. The literature had already started to build this connection even though it went often unnoticed. The BT model is equivalent to the maximization of an acyclic binary relation (Aleskerov et al., 2007, Theorem 4.1). The BAPU model is instead characterized by a stochastic choice function satisfying WST (Fudenberg et al., 2014).[4] The reader may not seen an immediate bridge between the two models. However, the connection will be made explicit in Section 3.4.2 via the completion of the results of Fishburn (1973). For instance, limited to the above mentioned, Fishburn (1973) showed that WST is the stochastic analogue of acyclicity.[5]

Our main results are twofold. First, we show that ε and η are metrics if and only

[3]The assumption of strict positivity of η is assumed to provide a connection to Fechnerian models as described in Section 3.4.1, however it is not necessary for the main results in Section 3.3 to hold.

[4]The result is an immediate corollary of (Fudenberg et al., 2014, Proposition 8). The reader may note that our model enriched with η is equivalent to the Item Invariant model of Fudenberg et al. (2014) restricted to binary sets. The restriction of c_A to $\eta(A)$ in binary sets is without loss of generality.

[5]More precisely, the property is called Acyclic Stochastic Transitivity [AST], which is equivalent to WST under the assumption of antisymmetry that will be described at the end of Section 3.2, and if all binary sets are observed.

if deterministic choices are outcomes of the maximization of a transitive binary relation and stochastic choices satisfy MST (Theorem 1 and 2). Second, ε and η are not metrics (the triangle inequality is violated) at particular triples of alternatives if and only if deterministic choices are outcomes of the maximization of a binary relation that satisfies a new property called Lower Negative Transitivity, and stochastic choices violate SST at the considered triples (Theorem 3 and 4). These results together provide a complete characterization of all notions of transitivity.

3.1.1 Related Literature

This Chapter is related to the deterministic choice literature on Threshold models: such authors include, among many, Luce (1956), Fishburn (1970) and, more recently, Frick (2016) and Dziewulski (2018). A comprehensive survey is provided by Aleskerov et al. (2007). Importantly, Theorem 1 has been proved using a contradiction argument by Nakamura (2002). In the conclusion of the paper, the author wrote: "It remains an open problem to give a constructive proof, which may also answer the question of whether arbitrary posets have quasi-metric threshold representations." We address part of the author's conclusion by providing a constructive proof for finite posets.

Regarding the stochastic choice literature, this Chapter relates mainly to Fudenberg et al. (2015), who characterize Additive Perturbed Utility models (APUs).[6] However, these models have also been studied by, among many, Machina (1985) and Mattsson & Weibull (2002) and they are the foundation of rational inattention literature (Matejka & McKay, 2015). Theorem 2 is new, albeit connected with a recent paper by He & Natenzon (2018).

Finally, this Chapter contributes to the literature on the connection between deterministic and stochastic choice theory. This topic requires a brief comment. The link between the two environments is far from straightforward. It seems natural to think that the only difference between deterministic and stochastic choice functions is the presence of probabilities, and that the former is a degenerate case of the latter

[6] In the published version, Fudenberg et al. (2015) only focus on APUs. Our model would be the restriction of theirs on binary sets if $\eta(x,y) = 1$ for all x,y. However, in the unpublished version, Fudenberg et al. (2014) characterize an extension of APUs, called Item Invariant APUs. Our model is a restriction of Item Invariant APUs on binary sets.

as in (Fudenberg et al., 2015, Proposition 6), in several results by Dasgupta & Pattanaik (2007) and as stated among others by Kalai et al. (2002), Gul & Pesendorfer (2006), and Cerreia-Vioglio et al. (2019). However, this is not the case as shown by Fishburn (1978) and Ok & Tserenjigmid (2019). The key distinction is related to what is considered as a *relevant comparison*. In the deterministic case, the only comparison that matters is the worst possible comparison.[7] Conversely, in the stochastic case, all comparisons matter as it is clear from the Beethoven/Debussy example (Debreu, 1960) and the Red/Blue Bus example (McFadden, 1974). In this Chapter, we focus on binary sets where the two interpretations are clearly equivalent and we complete the connection between the two environments initiated by Fishburn (1973). A generalization of our results to sets with higher cardinality is still, to the best of our knowledge, an unsolved problem.[8]

3.2 Preliminaries

Let A be a finite set of alternatives, $\mathcal{A}_2$ the set of all binary subsets of A and $\mathcal{A}_{1,2}$ the set of all singletons and binary subsets of A. A binary (set-valued) choice function is a mapping $c : \mathcal{A}_2 \to \mathcal{A}_{1,2}$ with $c(B) \subseteq B$ for all $B \in \mathcal{A}_2$. The following binary relation is defined starting from a primitive c:

$$x \succ_c y \Leftrightarrow x \in c(x,y) \ \& \ y \notin c(x,y)$$

The binary relation $\succ_c$ is asymmetric, and so irreflexive. It is also possibly incomplete. If $x,y \in c(x,y)$ then $x \not\succ_c y$ and $y \not\succ_c x$. This binary relation over pairs is foundational in the literature of deterministic choice theory, e.g. (Sen, 1971), (Arrow, 1959).

A binary stochastic choice rule is a mapping $p : A \times \mathcal{A}_2 \to [0,1]$ such that $p(x,A) + p(y,A) = 1$ for all $x,y \in A \in \mathcal{A}_2$ and $p(z,A) = 0$ for all $z \notin A$. With a common abuse of notation, we write $p(x,y)$ to denote the probability that x is chosen from $\{x,y\}$, and $p(\{x,y\})$ to denote the entire probability distribution.

[7] In other words, if there exists an element $\mathbf{x}$ that is noticeably better than the element $\mathbf{y}$, then $\mathbf{y}$ is never chosen when $\mathbf{x}$ is available, regardless of the other alternatives available.

[8] We see Fudenberg et al. (2014) and Fosgerau et al. (2017) as possible starting point for this problem.

Our results are based on two specific models of "just noticeable difference". The Binary Threshold models and Binary Additive Perturbed Utility models are defined as follows:

Definition 1. *Let $u : A \to \Re$ be a utility function and $\varepsilon : A \times A \to \Re^{++}$ be a threshold function. We say (u,ε) is a Binary Threshold Representation [BT] of a choice function c if for all $x,y \in A$:*

$$c(\{x,y\}) = \{x \in \{x,y\} \mid \nexists\, y \in \{x,y\} \mid u(y) - u(x) > \varepsilon(x,y)\} \tag{3.1}$$

Given the above definition, the following holds: $x \succ_c y$ if and only if $u(x) - u(y) > \varepsilon(x,y)$, and $[x \not\succ_c y \wedge y \not\succ_c x]$ if and only if $|u(x) - u(y)| \leq \varepsilon(x,y)$.

Definition 2. *Let $u : A \to \Re$ be a utility function, c be C^1 on $(0,1)$, a strictly convex function with $\lim\limits_{p\to 0} c'(p) = -\infty$ and $\eta : A \times A \to \Re^{++}$. We say (u,c,η) is a Binary Additive Perturbed Utility Representation [BAPU] for a stochastic choice rule p if for all $x,y \in A$:*

$$p(\{x,y\}) = \underset{p \in \Delta(\{x,y\})}{\operatorname{argmax}} \sum_{z \in \{x,y\}} u(z)p(z) - \eta(x,y) \cdot c(p(z)) \tag{3.2}$$

The existence of the above representations is characterized by restrictions on $\succeq_c$ and p. In particular, a choice function c has a BT representation if and only if $\succ_c$ is acyclic - Aleskerov et al. (2007).[9] A stochastic choice rule p has a BAPU representation if and only if p satisfies Weak Stochastic Transitivity [WST].[10]

The definition of BAPU needs two brief comments. First, in order to coherently match our analysis of Section 3.4.2 related to Fishburn (1973), we assume that $u(x) \neq u(y)$ for all $x,y \in A$ with $x \neq y$. This assumption does not modify the generality of the model but simply rules out the case of $p(x,y) = 0.5$. Second, the constraint $\lim\limits_{p\to 0} c'(p) = -\infty$ guarantees p to be non-degenerate. This assumption allows us to consider BAPU as equivalent to Fechnerian models. Similarly, we assume $\eta(x,y)$ to be strictly positive for all $x \neq y$. This constraint will have the effect of allowing ε to

[9] A binary relation $\succ_c$ is acyclic if for any integer k, $x_1 \succ_c x_2 \cdots \succ_c x_k$ implies $x_1 \neq x_k$.

[10] A stochastic choice rule p satisfies Weak Stochastic Transitivity if for all $x,y,z \in A$:

$$p(x,y) \geq 0.5 \;\;\&\;\; p(y,z) \geq 0.5 \quad \Rightarrow \quad p(x,z) \geq 0.5$$

be a quasi-metric[11], namely two different alternatives may have zero distance, while η will be a metric, namely none of the alternatives have $\eta(x,y) = 0$. The reader may note that the results hold if this constraint is relaxed in both directions, however in stochastic choice degenerate probabilities would be introduced.

3.3 Main results

This section contains the main results of the Chapter. The results rely **only** on $\varepsilon : A \times A \to \Re^{+}$ and $\eta : A \times A \to \Re^{++}$ satisfying or violating the triangle inequality.

Theorem 1. *A choice function c has a BT representation where ε is a quasi-metric if and only if the associated binary relation $\succ_c$ is transitive.*

Proof. The proof is constructive and presented in Appendix C.2.1. Here we present a description of the main steps of the proof. Firstly, we define the utility of an alternative x as the number of elements that are not strictly preferred to x. Secondly, we construct a weighted directed graph where the weights are defined as the difference in utility between the alternatives. We then show that the minimum weighted path $\delta(x,y)$ constitutes a quasi-metric. The final step involves scaling down $\delta(x,y)$ using a parameter $\gamma \in [0,1)$. In fact, note that for any $\gamma \in [0,1)$, $u(x) - u(y) > \gamma \cdot \delta(x,y)$ guarantees that $\succ$ is represented. Hence, the remaining challenge is to find γ such that for all $x,y \in X$ such that $x \not\succ y$ and $y \not\succ x$, $u(x) - u(y) \leq \gamma \cdot \delta(x,y)$. The problem of finding γ^* is shown to be the solution of the following maximization problem:

$$\underset{x,y \in A:[x \not\succ y \wedge y \not\succ x]}{\operatorname{argmax}} \frac{u(x) - u(y)}{\delta(x,y)}$$

$\square$

The application of the triangle inequality is confirmed by the result on Additive Perturbed Utility models. The restriction on p, as in He & Natenzon (2018), is Moderate Stochastic Transitivity*.

[11]The assumptions on the function ε are relaxed so as to be a quasi-metric, see Monjardet (1980). This relaxation does not change the nature of the result. A quasi-metric is a function $d : A \times A \to [0,\infty)$ that satisfies the following axioms: (1) $d(x,x) = 0$ "minimality"; (2) $d(x,y) = d(y,x)$ "symmetry"; (3) $d(x,z) \leq d(x,y) + d(y,z)$ for all $x,y,z \in A$ "triangle inequality". To be a metric, d has to satisfy also $d(x,y) = 0 \Leftrightarrow x = y$ "identity of indiscernibles".

Definition 3. *A stochastic choice rule p satisfies Moderate Stochastic Transitivity* [MST*] if for all $x,y,z \in A$ either of the following holds:*

(i) $p(x,y) > 0.5 \;\;\&\;\; p(y,z) > 0.5 \quad \Rightarrow \quad p(x,z) > \min[p(x,y), p(y,z)]$

(ii) $p(x,z) = p(x,y) = p(y,z)$

The reader may note that this property allows for the situation: $p(x,y) = p(x,z) > p(y,z)$; but it rules out the situation: $p(x,y) > p(x,z) = p(y,z)$.

Theorem 2. *A stochastic choice rule p has a BAPU representation where η is a metric if and only if it satisfies MST*.*

Proof. See Appendix C.2.2. □

For the next theorems we need to introduce a new deterministic property. We call this property Lower Negative Transitivity. First, we state some preliminaries. Suppose a finite set of alternatives A is weakly ordered[12] by a binary relation $\gg$. However, the decision maker has imperfect discrimination and does not observe the weak order $\gg$. Instead, he has a second binary relation $\succ$ that reveals if an alternative is "surely" better than another. We say that $\gg$ preserves $\succ$ if $x \succ y$ implies $x \gg y$ for all $x,y \in A$.[13] Dual to the weak order $\gg$ is a complete preorder $\trianglerighteq$.[14]

Definition 4. *A binary relation $\succ$ on a weakly ordered set $(A, \gg)$, preserved by $\gg$, satisfies Lower Negative Transitivity [LNT] at x,y,z if $x \trianglerighteq y \trianglerighteq z$ and $x \succ z$ implies either $x \succ y$ or $y \succ z$.*

The definition requires two short comments. First, notice that we rule out the possibility that $x \trianglerighteq y$ and $y \succ x$. Hence, $x \trianglerighteq y \trianglerighteq z$ implies $z \not\succ x$, $y \not\succ x$ and $z \not\succ y$. Second, if $x \succ z$ then we interpret this as x being "surely" better than z. The requirement that either $x \succ y$ or $y \succ z$ can be interpreted as follows: when the decision maker can discriminate between x,z then he can also discriminate between either x,y or y,z. The consequent result for Binary Threshold models is the following:

[12] A weak order $\gg$ is an asymmetric and negatively transitive binary relation.

[13] This approach has been studied by, among many, Fishburn (1999), Aleskerov et al. (2007), Ok & Nishimura (2018).

[14] A complete preorder $\trianglerighteq$ is a reflexive, complete and transitive binary relation. Due to the finiteness of A, such a binary relation has a numerical representation.

Theorem 3. *Let a choice function c have a BT representation. There exists a function ϵ, such that for all $x,y,z \in A$ with $u(x) \geq u(y) \geq u(z)$, $\varepsilon(x,z) \geq \varepsilon(x,y) + \varepsilon(y,z)$ if and only if the associated binary relation $\succ_c$ satisfies LNT at x,y,z.*

Proof. See Appendix C.2.3. □

This result connects violations of the triangle inequality with a new notion of transitivity. We can provide the same result using Additive Pertubed Utility models and a restriction on p called Strong Stochastic Transitivity*.

Definition 5. *A stochastic choice rule p satisfies Strong Stochastic Transitivity* [SST*] if for all $x,y,z \in A$ both of the following hold:*

(i) $p(x,y) > 0.5$ & $p(y,z) > 0.5 \quad \Rightarrow \quad p(x,z) \geq \max[p(x,y), p(y,z)]$

(ii) It cannot be that $p(x,z) = p(x,y) = p(y,z)$

Theorem 4. *Let a stochastic choice rule p have a BAPU representation. There exists a function η, such that for all $x,y,z \in A$ with $u(x) > u(y) > u(z)$; $\eta(x,z) \geq \eta(x,y) + \eta(y,z)$ if and only if p violates SST* at $x,y,z \in A$.*

Proof. The proof is constructive and presented in Appendix C.2.4. Here, we present the main steps of the sufficiency part which is the more involving. As in Theorem 1, we first assign a utility to each alternative simply using the ranking created by $p(x,y)$ that exists since WST is satisfied. Secondly, we rank couples of alternatives $\{x,y\}$ using again $p(x,y)$. Each couple now is assigned a strictly positive number $f(l)$ that is a function of their position l in the ranking. The number $f(l)$ will be then multiplied by the difference in utility to form $\eta(x,y) = f(l)|u(x) - u(y)|$. The remaining and more involving passage is to show that $f(l)$ is the solution of a particular difference equation, that when the initial condition is defined as $f(1) = 1$ and $f(2) = 2$ becomes:

$$f(l+1) = \frac{1 + 2(n-2) + f(l)}{1 + (n-2)}$$

The remaining step consists in constructing a continuous and differentiable cost function $c(p)$. This step is rather simple and described in the proof of Theorem 2 in Appendix C.2.2.

This constructive proof generalize one step of the proof of **?**. The authors construct one particular $f(l)$ that turns out to satisfy a similar version of the above difference equation that guarantee that triangle is satisfied instead of violated. $\square$

Theorem 4 requires a brief comment. The restriction on the stochastic choice rule p is not a general violation of SST*. To illustrate this, suppose that for all triples apart from x, y, z the stochastic choice rule p satisfies SST*. Then, since this latter is defined for all triples, we generically say p violates SST*. Theorem 4 states that in this case we can construct a function η that violates the triangle inequality exactly at those triples where SST* is violated, while it satisfies the triangle inequality in all other triples. This is a fundamental difference from Theorem 2, which states that there exists a η that satisfies the triangle inequality for all triples. A similar reasoning connects Theorem 1 and 3. Hence, the reader may refer to LNT and violation of SST* as local properties.

3.3.1 Summary

Theorems 2 and 4 are summarized in Figure 3.1. A similar diagram can be drawn for deterministic choice using properties of $\succeq_c$. The choices of $p(x,y)$ and $p(y,z)$ are arbitrary. The axis describes $p(x,z)$:

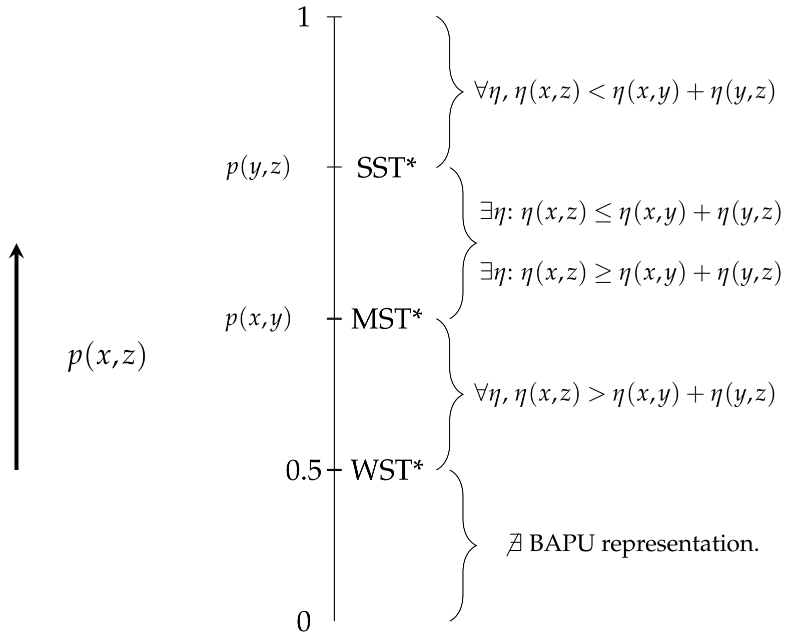


Figure 3.1. **Characterization of stochastic transitivity using triangle inequality.**

Let's show the implications of the results in view of the Paris-Rome example. Assume $u(\mathbf{P+}) = 2.1$, $u(\mathbf{P}) = 2$ and $u(\mathbf{R}) = 1$, and $\eta(\mathbf{P+},\mathbf{P}) = 0$, $\eta(\mathbf{P},\mathbf{R}) = \eta(\mathbf{P+},\mathbf{R}) = 1$. Assume also that the cost function is $c(p) = p \log p$ (Shannon Entropy). The resulting stochastic choice function is $p(\mathbf{P+},\mathbf{P}) = 1$, $p(\mathbf{P+},\mathbf{P}) = 0.75$, $p(\mathbf{P},\mathbf{R}) = 0.73$. Note that MST* is satisfied and η satisfies triangle inequality (note also that η is non-negative in this example and the result still holds but we observe degenerate probabilities). If we perturb $\eta(\mathbf{P+},\mathbf{R}) \pm \varepsilon$ with $\varepsilon > 0$ small, we either violate or strictly satisfy triangle inequality. However, $p(\mathbf{P+},\mathbf{R})$ is either 0.74 or 0.76, hence MST* is still satisfied. This shows that if the stochastic choice function satisfy MST* but violates SST*, we can construct η such that triangle inequality is either satisfied or violated as discussed in Theorem 2 and 4.

3.4 Related Literature

3.4.1 Fechnerian Model

Our results create a connection with a series of stochastic binary models, among which the most famous is the Fechnerian model. This model is defined as follows, given a utility function u and a strictly increasing function F:

$$p(x,y) = F[u(x) - u(y)]$$

The model has been axiomatized firstly by Debreu (1958). More recently, Fudenberg et al. (2015) proved an equivalence result with Additive Perturbed Utility models.

Proposition 1. *A stochastic choice rule p has a BAPU representation with $\eta(x,y) = 1$ for all $x, y \in A$ if and only if it has a Fechnerian representation.*[15]

A more general model, based on a distance function d, has been proposed by He & Natenzon (2018). The model, which we refer to as "HN representation", is as follows:

$$p(x,y) = F\left[\frac{u(x) - u(y)}{d(x,y)}\right]$$

The authors prove that this model is completely characterized by MST*. Therefore, the next corollary immediately follows:

Corollary 1. *A stochastic choice rule p has a BAPU representation where η is a metric if and only if it has a HN representation.*

3.4.2 Fishburn (1973)

Our analysis of BT and BAPU representations suggests a connection between deterministic and stochastic notions of transitivity. This problem has been firstly studied by Fishburn (1973). We build on his framework to complete the analysis and provide the reader with a full picture of this connection. If we let the stochastic choice rule p be primitive, we construct the following binary relation for any parameter $\mu \in [0.5, 1)$:

$$\succeq_\mu = \{(x,y) \in A \times A : p(x,y) \geq \mu\}$$

Note that the binary relation $\succeq_\mu$ is symmetric for some $x, y \in A$ if and only if

[15] This result is a restatement of Proposition 1 - Fudenberg et al. (2015).

$p(x,y) = 0.5$. Whenever $p(x,y) > 0.5$, for all $\mu \in [0.5,1)$ we can focus on the asymmetric part of $\succeq_\mu$, which we denote as $\succ_\mu$.

It is important to note that $\succ_c$ has not been assumed to be negatively transitive in Section 3.2. Hence, its symmetric part may describe indifference and most probably incompleteness. On the other hand, $\sim_\mu$ covers only an indifference relation, or equivalently a case where $u(x) = u(y)$. Similar reasoning can be generalized to cases of equalities within p as noticed by Fishburn (1973). When equalities are ruled out, or $p(x,y) \neq p(w,z)$ for all $x,y,z,w \in X$, the following properties are equivalent to the ones described in previous sections: AST $\Leftrightarrow$ WST, PST $\Leftrightarrow$ MST* and SST $\Leftrightarrow$ SST*. When one discards the measure zero set of equalities in p a clear bridge between the stochastic and deterministic worlds on binary sets arises.

Definition 6. *A stochastic choice rule p satisfies Acyclic Stochastic Transitivity [AST] if for any integer k and $x_1, x_2, \ldots, x_k \in A$:*

$$[p(a_1,a_2) > 0.5 \;\; \& \;\; p(a_2,a_3) > 0.5 \;\; \& \;\; \ldots \;\; \& \;\; p(a_{m-1},a_m) > 0.5] \;\; \Rightarrow \;\; p(a_1,a_m) \geq 0.5$$

Definition 7. *A stochastic choice rule p satisfies Partial Stochastic Transitivity [PST] if for all $x,y,z \in A$:*

$$p(x,y) > 0.5 \;\; \& \;\; p(y,z) > 0.5 \quad \Rightarrow \quad p(x,z) \geq \min[p(x,y), p(y,z)]$$

Definition 8. *A stochastic choice rule p satisfies Strong Stochastic Transitivity [SST] if for all $x,y,z \in A$:*

$$p(x,y) > 0.5 \;\; \& \;\; p(y,z) > 0.5 \quad \Rightarrow \quad p(x,z) \geq \max[p(x,y), p(y,z)]$$

Equivalence Results

AST and PST have been already connected to deterministic properties by Fishburn (1973). In particular, he proved that for all $\mu \in [0.5,1)$, $\succ_\mu$ is acyclic if and only if p satisfies AST; and $\succ_\mu$ is transitive if and only if p satisfies PST. No equivalence result has been provided for SST; in other words, there is no deterministic version of transitivity that has been shown to provide SST for all $\mu \in [0.5, 1)$. Some reasonings suggest that negative transitivity[16] can play a role; however, it has almost no explanatory power in stochastic terms, as was already noticed by Fishburn (1973). In particular, commenting on his Lemma 5b, he wrote about negative transitivity: "... when NST [Negative Stochastic Transitivity][17] holds, it requires a fair number of equalities within p. This may be viewed as further evidence of the general inapplicability of NST."

Our proposal to solve this problem is related to the concept of imperfect discrimination, and it has been anticipated with the introduction of LNT. We assume the existence of what Ok & Nishimura (2018) refer to as "preference structure". Say that $(A, \gg)$ is a weakly ordered set and $\succ$ is a binary relation that is preserved by $\gg$ as defined in Section 3.3. The dual of $\gg$ is again denoted as $\trianglerighteq$. We define one new property called Higher Negative Transitivity.

Definition 9. *A binary relation $\succ$ on a weakly ordered set $(A, \gg)$, preserved by $\gg$, satisfies Higher Negative Transitivity [HNT] if for all $x, y, z \in A$ such that $x \trianglerighteq y \trianglerighteq z$ we have: (1) $x \succ y$ implies either $x \succ z$ or $z \succ y$ and (2) $y \succ z$ implies either $x \succ z$ or $y \succ x$.*

This property, together with LNT, enables us to provide a characterization of SST in terms of deterministic properties, and to complete the analysis initiated by Fishburn (1973). Note that the two properties are defined differently. HNT is defined for all $x, y, z \in A$, while LNT is defined locally. The reasoning behind this definitions is contained in the following propositions.

Proposition 2. *A stochastic choice rule p either violates SST or satisfies it with equality at $x, y, z \in A$ if and only if in the same alternatives $\succ_\mu$ satisfies LNT for all $\mu \in [0.5, 1)$.*

Proof. Suppose $\succ_\mu$ violates LNT for some μ; then it must be that $p(x, z) > \mu$ while $p(x, y) < \mu$ and $p(y, z) < \mu$; but then SST is satisfied with inequality. Suppose SST is satisfied with inequality; then set $\mu = p(x, z) - \varepsilon$ and LNT is violated. $\square$

[16]A binary relation $\succ$ is negatively transitive if for all $x, y, z \in A$: $x \not\succ y$ and $y \not\succ z$ implies $x \not\succ z$. It can be equivalently stated as: $x \succ z$ implies either $x \succ y$ or $y \succ z$.

[17]Negative Stochastic Transitivity was defined as: $p(x, z) > 0.5 \Rightarrow \max[p(x, y), p(y, z)] \geq p(x, z)$. It was shown to be equivalent to $\succ$ satisfying *Negative Transitivity* [NT]. Hence, for all $x, y, z \in A$, if $x \not\succ y$ and $y \not\succ z$ then $x \not\succ z$.

Proposition 3. *A stochastic choice rule p satisfies SST if and only if $\succ_\mu$ satisfies HNT for all $\mu \in [0.5, 1)$.*

Proof. Suppose $\succ_\mu$ violates HNT, then either $p(x,y)$ or $p(y,z)$ are strictly greater than $p(x,z)$ violating SST. Suppose SST is violated, then set

$$\mu = \max[p(x,y), p(y,z)] - \varepsilon$$

for a small $\epsilon > 0$, and HNT is violated. □

Before summarizing the results in an implication diagram, we describe in a remark what are the connections between transitivity and the two new properties LNT and HNT.

Remark 2. Since transitivity is equivalent to PST, it is implied by HNT, which in fact is equivalent to SST. Instead, LNT and transitivity are not connected. For instance, $x \succ y \succ z$ violates transitivity, but not LNT; $x \succ z$ satisfies transitivity but violates LNT. Similarly, if p satisfies only AST, then $\succ$ satisfies LNT but not transitivity; while if p satisfies SST with inequality, then $\succ$ satisfies transitivity but not LNT.

We are now ready to complete part of the implication diagram initiated by Fishburn (1973). We will denote the special case where SST is satisfied with equality as SST$^{\approx}$, and the case where it is violated with inequality or satisfied with equality as $\overline{\text{SST}}$. Hence, it must be that SST$^{\approx}$ = SST $\wedge$ $\overline{\text{SST}}$. All the results of this section are summarized in the diagram below, and show a perfect symmetry between the deterministic and stochastic versions of transitivity.

$$\begin{tikzcd}
\mathrm{NT} \arrow[rr, mapsto] \arrow[d, leftrightarrow] & & \mathrm{HNT} \arrow[rr, mapsto] \arrow[dr, |-|] & & \mathrm{T} \arrow[rr, mapsto] \arrow[dl, |-|] & & \mathrm{A} \arrow[d, leftrightarrow] \\
\mathrm{HNT} \wedge \mathrm{LNT} \arrow[rrr, mapsto] & & & \mathrm{LNT} \arrow[rrr, mapsto] \arrow[urrr, |->] & & & \mathrm{HNT} \vee \mathrm{LNT}
\end{tikzcd}$$

$$\begin{tikzcd}
\mathrm{SST}^{\approx} \arrow[rr, mapsto] \arrow[d, leftrightarrow] & & \mathrm{SST} \arrow[rr, mapsto] \arrow[dr, |-|] & & \mathrm{PST} \arrow[rr, mapsto] \arrow[dl, |-|] & & \mathrm{AST} \arrow[d, leftrightarrow] \\
\mathrm{SST} \wedge \overline{\mathrm{SST}} \arrow[rrr, mapsto] & & & \overline{\mathrm{SST}} \arrow[rrr, mapsto] \arrow[urrr, |->] & & & \mathrm{SST} \vee \overline{\mathrm{SST}}
\end{tikzcd}$$

Figure 3.2. **Implication Diagram for the results of Section 3.4.2.**

3.5 Conclusion

This Chapter provides a new characterization of stochastic transitivity related to the well-known concept of imperfect discrimination. As the main message, we show that in Fechnerian models, triangle inequality and transitivity are closely connected. This notion allows us to organize a wide range of stochastic models in accordance with a very general version of Fechnerian models. Furthermore, given that all the proofs are constructive, we provide simple algorithms to construct both the utility function and imperfect discrimination parameter.

Appendix A

Appendix to Chapter 1

A.1 Minor Results and Remaining proofs

A.1.1 Proof of Proposition 2

By transitivity, completeness of R and the finiteness of X; we can make use of a result from Krantz et al. (1971): there exists a real-valued function ϕ on X such that for all $x, y \in X$; xRy if and only if $\phi(x) \geq \phi(y)$.

A corollary of this result goes as follows: let $\phi : X \to R^{n-1}$, where $|X| = n$, be a vector valued function and $\phi(x)_z$ be the value assigned to x when compared to z. Then, by the previous result, $\phi(x)_z = \phi(x)_y$ for all $y, z \neq x$. The proof is trivial. Suppose the above is false; then we may have $\phi(x)_y > \phi(y)_z > \phi(z)_x$ violating transitivity.

Given two generic elements x, y we can partition the collection of observations in eight sets with the following cardinalities: C_{xy}, C_{yx} have already been defined; $C_{x,-y} = |\{S \in D : x = C(S), y \notin S\}|$ and similarly $C_{y,-x}$; $B = B_{xy} = B_{yx} = |\{S \in D : z = C(S); x, y \in S\}|$; $D_{xy} = |\{S \in D : z = C(S), x \in S, y \notin S\}|$ and similarly D_{yx}; $E = E_{xy} = E_{yx} = |\{S \in D : z = C(S); x, y \notin S\}|$.

Let's first focus on B and E. On these collection of observations, Neutrality implies $xI^{C_D}y$. Suppose $u(x) \geq u(y)$, then $C_{x,-y} \geq C_{y,-x}$. By induction, suppose $C_{x,-y} + C_{y,-x} = 1$ such that $x = C(S)$, then by Choice non-negativeness $xR^{C_D}y$. Suppose the hypothesis holds for $C_{x,-y} + C_{y,-x} = n$, and take $C_{x,-y} + C_{y,-x} = n + 1$ (note that in both cases $C_{x,-y} \geq C_{y,-x}$). Suppose by contradiction that $yP^{C_D}x$. Then, if we remove one observation where x is chosen, Choice non-negativeness is violated. Hence, $xR^{C_D}y$. Similarly, $C_{xy} \geq C_{yx}$ implies $xR^{C_D}y$ by Informational Responsiveness, Choice non-negativeness and Neutrality. Note that $u(x) \geq u(y)$ implies $C_{x,-y} \geq C_{y,-x}$ and $C_{xy} \geq C_{yx}$ doesn't hold for a generic domain D. However, it holds for a homogeneous domain.

To complete the proof we need to extend the argument to D_{xy} and D_{yx}. However, note that $u(x) \geq u(y)$ implies $D_{yx} \geq D_{xy}$ and there are no constraints on how such observations should influence the ranking between x, y since a third element is chosen. Hence, a method that associate a positive value to the observations of the type D_{xy}, D_{yx} could led to $u(x) > u(y)$ and $yP^{C}x$. However, by the corollary of Krantz et al. (1971) result, which is based on Transitivity, we can focus on $D_x = \sum_{z \neq x} |\{S \in D : z = C(S), x \in S\}|$ instead of D_{xy}. In other words, the value assigned by a method to the observation $z = C(S)$, $x \in S$, $y \notin S$ must be equal to the one assigned to $y = C(S)$, $x \in S$, $z \notin S$.; otherwise this could potentially lead to cycles. Hence, suppose by contradiction that $u(x) > u(y)$ and $yR^{C_D}x$. It must be that the value attached to observations in D_x is positive, since $D_y > D_x$. However, we proved that $xP^{C_D}y$ over the collections of observations with cardinalities $C_{x,-y}$, $C_{y,-x}$,C_{xy}, C_{yx}, B, E. Suppose we add an observation of the type $x = C(S)$, $y \in S$. Clearly, D_y increase by a positive value. However, since we assumed $yR^{C_D}x$ then Choice non-negativeness is violated. In words, these axioms allow a method to attach a positive value to observations of the type D_x, however this value must be smaller than the value attached to observations of the type C_{xy} as clear from the following example.

Example 1. *Take $X = \{x, y\}$, $u(x) > u(y)$ and a method g such that:*

$$xR^{C}_{g}y \Leftrightarrow H_x \geq H_y$$

where $H_x = a \cdot C_x + b \cdot D_x$ with $b = 2 > a = 1$.

$$\left[\begin{array}{c|c} S & \{x,y\} \\ C(S) & (5,5) \end{array}\right] \qquad \left[\begin{array}{c|c} S & \{x,y\} \\ C(S) & (6,5) \end{array}\right]$$

From the first dataset we infer $H_x = 15$ and $H_y = 15$ and so $xI^{C_D}y$; from the second dataset we infer $H_x = 16$ and $H_y = 17$ and so $yP^{C_D}x$ violating Choice non-negativeness (and in this case also Informational Responsiveness).

A.1.2 Some auxiliary facts about Counting Procedures.

The reader may have observed that when the restrictions of *Theorem 2* apply, adding information implies removing information in the definition of (Strong) Informational Responsiveness. Hence, this latter could be, in principle, omitted by the definition.

Claim 1. *If a method g satisfies Choice non-negativeness then the second consequent in the definition of Informational Responsiveness can be omitted.*

Proof. Suppose by contradiction that if $xI^{C_D}y$ and $x = C(S)$ then $xR^{C_{D\setminus S}}y$. Then, if $xI^{C_{D\setminus S}}y$, we immediately contradict Informational Responsiveness in its first consequent. If $xP^{C_{D\setminus S}}y$, then Choice Non-Negativeness is violated.

More interesting than the trivial proof is an example that shows the independence of adding and removing data. Let $xP^{C_D}y$ when $|D| < 2$ and $xR^{C_D}y$ if and only if $C_x \geq C_y$ when $|D| \geq 2$. Moreover, **CC** holds for all other $z \neq x, y$. Notice that this method violates Neutrality and Choice non-negativeness. However, it satisfies Informational Responsiveness in its first consequent since it does vacuously when $|D| < 2$. However, it violates Informational Responsiveness in its second consequent. Let $x = C(S)$ and $y = c(T)$, then $xI^{C_D}y$. If S is removed then $xP^{C_D}y$ violating the second consequent. □

In *Section 1.6* we introduced Robustness as an appealing characteristic of welfare methods. We also claimed that it is redundant in proving *Theorem 2* since it is implied by the collection of other axioms. This result can be shown easily as a corollary of *Theorem 2*, using the transitivity property of the counting choice method.

Corollary 1 (Theorem 2). *Independence, Neutrality, Stability and Strong Informational Responsiveness imply Robustness*

Proof. Suppose, by contradiction, $xP^{C_D}zP^{C_D}y$ and $yR^{C_{D\setminus S}}x$. If $yP^{C_{D\setminus S}}x$ then Stability is contradicted. Suppose $yI^{C_{D\setminus S}}x$. If $z = C(S)$ then Independence is contradicted. If

$y = C(S)$ then Strong Informational Responsiveness is contradicted. If $x = C(S)$ then we exploit the transitivity property of $C_x \geq C_y$. In fact, by Independence $zP^{C_{D\setminus S}}y$; by Stability either $xI^{C_{D\setminus S}}z$ or $xP^{C_{D\setminus S}}z$, and in both cases transitivity is contradicted. □

A.1.3 Proof of Proposition 3

Similarly to the others, the proof is by induction on the cardinality of the domain.

If $|D| = 0$ then $xI^{C_D}y$ by Neutrality and Completeness. Suppose $|D| = 1$ if $x = C(S)$ then $xP^{C_D}y$ by Strong Positive Responsiveness. If $z = C(S)$ then $xI^{C_D}y$ by Independence.

Suppose $|D| = 2$ and $C_x = C_y$ then:

- if $x = C(x,y)$, $y = C(x,y)$ then $xI^{C_D}y$ by Neutrality.
- if $x = C(x,y)$, $y = C(y,z)$ then suppose $xP^{C_D}y$. If we add $z = C(x,z)$ we have $xP^{C_{D\cup\{x,z\}}}y$ by Independence. So, the possible results by Transitivity are $xP^{C_{D\cup\{x,z\}}}yP^{C_{D\cup\{x,z\}}}z$ or $zP^{C_{D\cup\{x,z\}}}xP^{C_{D\cup\{x,z\}}}y$ or $xI^{C_{D\cup\{x,z\}}}zP^{C_{D\cup\{x,z\}}}y$ or $xP^{C_{D\cup\{x,z\}}}zP^{C_{D\cup\{x,z\}}}y$. However, take the permutation $\pi(x) = y$, $\pi(y) = z$ and $\pi(z) = x$. The choice function is preserved while the binary relations are not. Hence, we have a contradiction, implying $xI^{C_D}y$.
- if $x = C(x,z)$, $y = C(y,w)$. Then, take $\pi(x) = y$, $\pi(y) = x$, $\pi(z) = w$ and $\pi(w) = z$ and by the same argument as before $xI^{C_D}y$.

Now, suppose $|D| = n$, assume the statement is true and take $|D \cup \{T\}| = n+1$. Suppose $C_x > C_y$ then if x or y are chosen from T, $xP^{C_{D\cup T}}y$ by Strong Positive Responsiveness and the inductive hypothesis. If z is chosen, $xP^{C_{D\cup T}}y$ by Independence and the inductive hypothesis.

Suppose $C_x = C_y$ and $z = C(T)$ then the results holds by Independence. If y or x is chosen then $xP^{C_D}y$ (or $yP^{C_D}x$) by the inductive hypothesis. Hence, suppose by contradiction that $yP^{C_{D\cup T}}x$, then there exists a set T such that x is chosen since $C_x = C_y$. Hence, by the inductive hypothesis, Strong Positive Responsiveness is violated since $xP^{C_D}y$.

A.1.4 Proof of Theorem 3

Let $|D| = 0$, then by Neutrality and Completeness $xI^{C_D}y$. Let $D = \{S\}$ and $z = C(S)$. If either $x \notin S$ or $y \notin S$ or both then $xI^{C_D}y$ by Connection. If $x, y \in S$ then by Neutrality $xI^{C_D}y$. Suppose that $x = C(S)$; if $y \notin S$ then $xI^{C_D}y$ by Connection. If $y \in S$ then $xP^{C_D}y$ by Informational Responsiveness. Take $D = \{S, T\}$. If $z = C(T)$ and $C_{xy} = C_{yx} = 0$ then $xI^{C_D}y$ by Connection and Neutrality; if $C_{xy} = 1$ then $xP^{C_D}y$ again by Connection and Neutrality. Let $x = C(T)$; if $C_{xy} = 2$ then $\neg yP^{C_D}x$ by Stability; suppose $xI^{C_D}y$ then it should be $yP^{C_{D \setminus T}}x$ by Informational Responsiveness, but this contradicts the premise; hence $xP^{C_D}y$. If $C_{xy} = C_{yx} = 1$ then $xI^{C_D}y$ by Stability and Completeness.

Suppose the result holds for $|D| = n$ and take $|D \cup T| = n + 1$. If $C_{xy} = C_{yx}$ and $x = C(T)$ then $x^{C_{D \cup T}}y$ by Stability and Completeness. If $z = C(T)$ then the results holds by either Connection or Neutrality. If $C_{xy} - C_{yx} = 1$ then $xP^{C_{D \cup T}}y$ by Informational Responsiveness and the inductive hypothesis. If $C_{xy} - C_{yx} > 1$ then $\neg yP^{C_{D \cup T}}x$ by Stability and the inductive hypothesis. Suppose $xI^{C_{D \cup T}}y$, then removing $x = C(T)$ it should be $yP^{C_D}x$ contradicting the inductive hypothesis, hence $xP^{C_{D \cup T}}y$.

A.2 Comments on Additivity and Continuity

A.2.1 Continuity

Our axioms are defined over an abstract setting. In this appendix section we analyse the problem from a different perspective. The rationale for this analysis is that Stability reminds closely Bolzano's theorem (which is the equivalent of intermediate value theorem around zero). However, to draw this connection we need to endow domain and codomain with a metric topology.

A method is a map from a set of choice functions to a set of binary relations. If we focus on **CC** we can alternatively see it as a map that assigns to any alternative a non-negative integer. In this case, we can endow the set of non-negative integers with the metric $|x - y|$ for all $x, y \in \mathcal{Z}$. Since a method maps to a set of binary relations, we then write $xP^{C}_{\mathbf{CC}}y$ if and only if $C_x - C_y > 0$ and $xI^{C}_{\mathbf{CC}}y$ if and only if $C_x - C_y = 0$.

It is less straightforward to endow the space of choice functions $\mathscr{C}$ with a metric. However, we can define a metric between two choice functions $C_i, C_j \in \mathscr{C}$ using the symmetric distance (Klamler, 2008): $d(C_i, C_j) = \sum_{S \in K} |\Delta(C_i(S), C_j(S))| =$ where $|\Delta(C_i(S), C_j(S))| = |(C_i(S) \cup C_j(S)) \setminus (C_i(S) \cap C_j(S))|$. Klamler (2008) proposes this distance restricted to a fix domain D. We generalize it allowing D to change while keeping the choice function C fixed. In this case we can just take the symmetric difference among domains. Hence, for all $D_1, D_2 \in \mathcal{D}$, $\Delta(D_1, D_2) = |(D_1 \cup D_2) \setminus (D_1 \cap D_2)|$. It is easy to verify that for a fixed domain $D \subseteq \mathcal{X}$ and for any three choice functions $C_i^D, C_j^D C_l^D \in \mathscr{C}(D)$:

$d(C_i^D, C_j^D) \geq 0$

$d(C_i^D, C_i^D) = 0$

$d(C_i^D, C_j^D) = d(C_j^D, C_i^D)$

$d(C_i^D, C_l^D) \leq d(C_i^D, C_j^D) + d(C_j^D, C_l^D)$.

$\min_{C_i^D \neq C_j^D} d(C_i^D, C_j^D) = 2$

On the other hand, if for a fixed choice function $C \in \mathscr{C}$, and for any three domains $D, D^*, D' \in \mathcal{D}$:

$d(C^D, C^{D^*}) \geq 0$

$d(C^D, C^D) = 0$

$d(C^D, C^{D^*}) = d(C^{D^*}, C^D)$

$d(C^D, C^{D'}) \leq d(C^D, C^{D^*}) + d(C^{D^*}, C^{D'})$ [1]

$min_{D \neq D^*} d(C^D, C^{D^*}) = 1$

A welfare method is now a function between metric spaces. We recall the general definition of ϵ, δ continuity:

Definition. *A function $f : R \to R$ is continuous at $p \in R$ if for any $\epsilon > 0$ there exists $\delta > 0$ s.t. for all $x \in R$ if $|p - x| < \delta$ then $|f(p) - f(x)| < \epsilon$.*

In this classical definition, we can set a δ as function of x and ϵ can be any number

[1] Let's take $D_1, D_2, D_3 \in \mathcal{X}$: $d(D_1, D_3) = |(D_1 \setminus D_3) \cup (D_3 \setminus D_1)|$. Given D_2 we have $D_1 \setminus D_3 = (D_1 \setminus (D_2 \cup D_3)) \cup ((D_1 \cap D_2) \setminus D_3)$. So, show that $(D_1 \setminus D_3) \cup (D_3 \setminus D_1) \subseteq (D_1 \setminus D_2) \cup (D_2 \setminus D_1) \cup (D_2 \setminus D_3) \cup (D_3 \setminus D_2)$:

$[(D_1 \setminus (D_2 \cup D_3)) \cup ((D_1 \cap D_2) \setminus D_3)] \cup [(D_3 \setminus (D_1 \cup D_2)) \cup ((D_2 \cap D_3) \setminus D_1)] \subseteq [(D_1 \setminus (D_2 \cup D_3)) \cup ((D_1 \cap D_3) \setminus D_2)] \cup [(D_2 \setminus (D_1 \cup D_3)) \cup ((D_1 \cap D_2) \setminus D_3)] \cup [(D_2 \setminus (D_1 \cup D_3)) \cup ((D_2 \cap D_3) \setminus D_1)] \cup [(D_3 \setminus (D_1 \cup D_2)) \cup ((D_1 \cap D_3) \setminus D_2)]$

In fact, note that $[(D_1 \setminus D_2) \cup (D_2 \setminus D_1) \cup (D_2 \setminus D_3) \cup (D_3 \setminus D_2)] \setminus [(D_1 \setminus D_3) \cup (D_3 \setminus D_1)] = ((D_1 \cap D_3) \setminus D_2) \cup (D_2 \setminus (D_1 \cup D_3)) \cup (D_2 \setminus (D_1 \cup D_3)) \cup ((D_1 \cap D_3) \setminus D_2)$

however small. In the discrete case, both δ and ϵ have some bounds. We rely on Johnsonbaugh (1998), who introduces an intermediate value theorem for integer-valued functions with a definition of continuity where $\delta = \epsilon = 1$:

Definition. *Let f an integer-valued function defined on the integers in $[m,n]$. Suppose (as the equivalent of a continuity assumption) that $|f(i) - f(i+1)| \leq 1$ for $m \leq i < n$. If $f(m)f(n) < 0$, then $f(x) = 0$ for some integer x in (m,n).*

For example, if we focus on **CC**, continuity is based on a "marginal change" δ in the domain which is $\min_{d>0} d(C_i, C_j) = 1$. Since the codomain is the set of non-negative integers with the metric $|x - y|$, the minimum change greater than zero is $\epsilon = 1$. Hence, the counting choice method, **CC** : $(\mathscr{C}(D), d) \to (\mathcal{Z}, |\cdot|)$, can be characterized using Neutrality, Strong Informational Responsiveness, Independence and (1-1)-Continuity.

In general, the definition of Continuity depends on the distance defined on the space of choice functions. This distance is primitive and, in fact, using Stability we assume the marginal change to be $\delta = |D| - |D \setminus \{S\}|$. However, the reader may imagine a different δ that changes with respect to the cardinality of the set involved in the marginal change. In principle, one could consider a bigger set more important than a smaller one. Once δ is defined, then ϵ will describe the "degree" of continuity of the function. For instance, if $\epsilon = 2$, then **CC** would be continuous among intervals of 2, while into an interval of 2 it could move discontinuously. Therefore, we can interpret the level of ϵ as a measure of robustness in the following way: if $xPz_1P \ldots Pz_\epsilon Py$ then $xP^{C_{D \setminus S}}y$.

A.2.2 Additivity

In this appendix section, we comment on a second framework that may be of interest: conjoint measurement. When we think about a function that ranks the alternatives, it could be of some interest to think of it as a function that extracts some information about the importance of each alternative from each set. This definition is in line with the literature of conjoint measurement where each set is considered as an attribute of the alternative importance. Krantz et al. (1971) shows that given a product space, there exists an additive conjoint representation only if R_g is rational, independent and solvable. We refer the reader to Krantz et al. (1971) for the result and definitions.

The reader may note that **CC** clearly admits an additive conjoint representation. An interesting exercise is to verify if other methods such as **MS** and **EIG** have such representation. The task is quite straightforward since the independent condition is necessary.

Definition. *A preference relation R_g is independent if $(x_i, x_{-i}) R_g (y_i, x_{-i})$ is equivalent to $x_i R_g y_i$ for all $i \in \{1, 2, \dots, n\}$.*

The conjoint representation over the product space can be described using a function F given a choice function C and a list of sets S. $F(C(S_i))_{S \in D}$ is $F : X^{|D|} \to \mathscr{R}^{|X|}$ while the function of a single element is $F_x : X^{|D|} \to \mathscr{R}$. Then if additivity holds, we would have $F_x(C(S_i))_{S \in D} = \phi_x(C(S_1)) + \cdots + \phi_x(C(S_n))$ for some function ϕ.

Our examples will make use of $x I_g^{C_D} y$ via the following trick: take a permutation π as defined in Chapter 1 such that $\pi(x) = y$, $\pi(y) = x$ and $\pi(z) = z$ for all $z \neq x, y$; then use Neutrality to provide the contradiction. The counterexamples that involve **MS** and **EIG** will make this trick clear.

Take the choice function: $C(x, y) = x$, $C(x, y, z) = y$, $C(x, z) = z$. The swaps solution here suggests $y P_{\mathbf{MS}}^{C_D} x$. By Neutrality the permutation π produces $x P_{\mathbf{MS}}^{C_D} y$. Let's now add $C(x, z) = z$ on both the choice functions. In the original choice function we still have $y P_{\mathbf{MS}}^{C_{D \cup \{x,z\}}} x$, but in the one after the permutation we have $x I_{\mathbf{MS}}^{C} y$. This suggests that the new set is helping y in some way. But then by Krantz's independence we should have $y P_{\mathbf{MS}}^{C_{\{x,z\}}} x$ only over $z = C(x, z)$ which is a contradiction. Hence, the preference relation resulted from **MS** is not independent into the choice function and so there exists no additive conjoint representation such that $F_x(C(S_i))_{S \in K} \geq F_y(C(S_i))_{S \in K} \Leftrightarrow x R_{MS} y$.

The same can be proved for the **EIG** method. Take the following choice function: $C(x, z) = x$, $C(w, y) = w$. Here, $x P_{\mathbf{EIG}}^{C_D} y$ and given a permutation π, by Neutrality, $y P_{\mathbf{EIG}}^{C_D} x$. Let's now add other two choices, namely $C(w, y) = w$ and $C(z, w, y) = z$. Here, in the first choice function we have $x P_{\mathbf{EIG}}^{C_{D \cup \{w,y\}, \{z,w,y\}}} y$, but in the second one we have $x I_{\mathbf{EIG}}^{C_{D \cup \{w,y\}, \{z,w,y\}}} y$. This suggests that the two new choices help x over y. But, **EIG** doesn't discriminate between two elements if they are never chosen; even though they are beaten a different number of times. Hence, $F_x(C(w,y) = w, C(z,w,y) = z) = F_y(C(w,y) = w, C(z,w,y) = z)$; which contradicts the independence condition.

Appendix B

Appendix to Chapter 2

B.1 Alternatives

We start presenting the rationale behind the parametrization of delayed payment plans and lotteries. A drawback of the experiments of Manzini et al. (2010) and Barberá & Neme (2017) was that some alternatives were dominated if the decision maker happened to be a discounter. This problem led to a very high number of rational subjects and therefore a low level of heterogeneity in preferences.

Following Agranov & Ortoleva (2017), we construct the MAIN alternatives with no clear domination. We run a pilot experiment in the game theory classes of the Queen Mary University of London to confirm the presence of heterogeneity in preferences.

B.1.1 Time

Table B.1. **List of Delayed Payment Plans**

ALTERNATIVES		MONTHS 0	3	6	9	12
ONE SHOT (OS)		**160**	**0**	**0**	**0**	**0**
	M(OS)	140	0	0	0	0
	Im(OS)	0	160	0	0	0
DECREASING (D)		**110**	**50**	**25**	**0**	**0**
	M(D)	100	40	10	0	0
	Im(D)	0	110	50	25	0
CONSTANT (K)		**50**	**50**	**50**	**50**	**0**
	M(K)	40	40	40	40	0
	Im(K)	0	50	50	50	50
INCREASING (I)		**0**	**15**	**40**	**170**	**0**
	M(I)	0	10	20	160	0
	Im(I)	0	0	15	40	170
NEUTRAL						
	Neu1	15	55	30	20	5
	Neu2	5	20	30	55	15

NOTE -- The amounts are described in Token. The exchange rate was fixed at 20:1 pounds. The confounding alternatives are divided in: (1) "M" or Monotonicity Dominated; (2) "Im" or Impatience Dominated. The first is obvious; the second regards sequence of payments with the same total summation but paid three months later.

Table B.1 presents the comprehensive list of delayed payment plans. We use a quasi-hyperbolic discounting model[1]: $u(x_0, \dots x_t) = x_0 + \beta \sum_{t=1,2,\dots} \gamma^t x_t$. This specification displays a present bias for $\beta < 1$. Among many Benhabib et al. (2010) find that in an experimental setting subjects display a significant present bias[2]. We set $\beta = 0.9$ such that the following preferences arise for different levels of γ:

[1]See Laibson (1997) or Phelps & Pollak (1968) for a review of quasi-hyperbolic discounting. Note the following difference in modelling discount utility:

- Exponential Discounting: $v(x_0, \dots x_t) = \sum_{t=0,1,2,\dots} \beta^t u(x_t)$;
- Quasi-Hyperbolic Discounting: $v(x_0, \dots x_t) = u(x_0) + \beta \sum_{t=1,2,\dots} \gamma^t u(x_t)$;
- Hyperbolic Discounting: $v(x_0, \dots x_t) = \sum_{t=0,1,2,\dots} [\prod_{t=0,1,2,\dots} \gamma(t)] u(x_t)$.

[2]Benhabib et al. (2010) use a slightly different specification with fixed cost to represent the present bias.

$$\begin{array}{c|c} \textbf{Discount rate} & \textbf{Preference} - (\succ) \end{array}$$
$$\left[\begin{array}{c|c} 0.985 < \gamma \leq 1 & \mathbf{I} \succ \mathbf{K} \succ \mathbf{D} \succ \mathbf{OS} \\ 0.982 < \gamma \leq 0.985 & \mathbf{I} \succ \mathbf{D} \succ \mathbf{K} \succ \mathbf{OS} \\ 0.974 < \gamma \leq 0.982 & \mathbf{D} \succ \mathbf{K} \succ \mathbf{I} \succ \mathbf{OS} \\ 0.965 < \gamma \leq 0.974 & \mathbf{D} \succ \mathbf{K} \succ \mathbf{OS} \succ \mathbf{I} \\ 0.926 < \gamma \leq 0.965 & \mathbf{D} \succ \mathbf{OS} \succ \mathbf{K} \succ \mathbf{I} \\ 0 < \gamma \leq 0.926 & \mathbf{OS} \succ \mathbf{D} \succ \mathbf{K} \succ \mathbf{I} \end{array}\right]$$

The construction is meant to penalise the two simple alternatives, **OS** and **K**. The former is either first or second best for $\gamma < 0.965$, which would imply an annual discount rate of about 0.65; while the latter is never a first-best alternative. This choice has been driven by the necessity of avoiding "simplicity seeking" heuristics as observed by Iyengar & Kamenica (2010). Conversely, the main difference between **D** and **I** payment plans is the present bias, which is a feature of quasi-hyperbolic discounting. In fact, without present bias we should observe **I** chosen by most individuals since, in an exponential model, it is the best element for any annual discount rate bigger than 0.974. The exponential model maintains the feature that **K** and **OS** are never the best alternatives. This reveals that any choice of these two payment plans in a set with all alternatives available is due to (i) heuristics; (ii) exceptionally low discount rate; (iii) negative or non-monotone time preferences that would need the more complex hyperbolic discounting model to be encountered.

B.1.2 Risk

Table B.2. **List of Lotteries**

ALTERNATIVES		TOKEN		PROBABILITIES		EV	SD
DEGENERATE (D)		**50**	**0**	**1**	**0**	50	0
	F1(D)	45	0	1	0	45	0
	F2(D)	50	0	0.95	0.05	47.5	10.9
	So(D)	51	0	0.95	0.05	48.45	11.12
SAFE (S)		**65**	**25**	**0.8**	**0.2**	57	16
	F1(S)	60	20	0.8	0.2	52	16
	F2(S)	65	25	0.75	0.25	55	17.32
	So(S)	70	15	0.7	0.3	53.5	25.2
FIFTY-FIFTY (50)		**95**	**20**	**0.5**	**0.5**	57.5	37.5
	Sim	100	0	0.5	0.5	50	50
	F2(50)	95	20	0.45	0.55	53.75	37.31
	So(50)	110	20	0.4	0.6	56	44.09
RISKY (R)		**300**	**5**	**0.2**	**0.8**	64	118
	F1(R)	250	5	0.2	0.8	54	98
	F2(R)	300	5	0.15	0.85	49.25	105.34
	So(R)	500	5	0.1	0.9	54.5	148.5

NOTES -- The amounts are described in Token. The exchange rate was fixed at 10:1 pounds. The confounding alternatives are divided in: (1) "F1" and "F2" - First Order Stochastically Dominated; (2) "So" - Second Order Stochastically Dominated; (3) Sim - SIMPLE. This latter was created to check for "simplicity seeking" heuristics connected to simple numbers.

Table B.2 presents the comprehensive list of lotteries. The parametrization follows a CRRA utility function with parameter γ such that the following preferences arise:

$$\begin{array}{cc} \textbf{Risk parameter} & \textbf{Preference} - (\succ) \\ \left[\begin{array}{c|c} 0 \leq \gamma < 0.1 & \mathbf{R} \succ \mathbf{50/50} \succ \mathbf{S} \succ \mathbf{D} \\ 0.1 \leq \gamma < 0.2 & \mathbf{S} \succ \mathbf{50/50} \succ \mathbf{R} \succ \mathbf{D} \\ 0.2 \leq \gamma < 0.6 & \mathbf{S} \succ \mathbf{50/50} \succ \mathbf{D} \succ \mathbf{R} \\ 0.6 \leq \gamma < 1.9 & \mathbf{S} \succ \mathbf{D} \succ \mathbf{50/50} \succ \mathbf{R} \\ \gamma \geq 1.9 & \mathbf{D} \succ \mathbf{S} \succ \mathbf{50/50} \succ \mathbf{R} \end{array}\right] \end{array}$$

Some evidence about estimation of risk aversion parameters in laboratory (Bombardini & Trebbi, 2012), (Harrison & Rutstrom, 2008), (Soltani et al., 2012) and on field (Kim & Lee, 2012) seem to suggest that in binary sets **S**, **50/50** and **R** should obtain a substantial amount of choices since $\gamma > 1$ has been rarely observed, especially in laboratory experiments. **R** has been designed to attract risk neutral individuals, while **50/50** has been slightly penalised against **S** due to the fact that it could be chosen using simplicity seeking heuristics (Iyengar & Kamenica, 2010). In general, the amount of token in the three non-degenerate lotteries has been set in order to avoid simple amounts such as (10, 100 or 1000) and make the calculation of expected values not straightforward. Finally, the **SIMPLE** lottery (100,0) with 50 percent probability has been created to test the robustness of the simplicity seeking argument. Notice also that Second Order Stochastically Dominated alternatives have a smaller mean than the MAIN alternatives. This was to guarantee a sufficient choice of MAIN alternatives as well as highlight subjects with risk-seeking behaviour.

B.2 Order of Questions

Another drawback of previous experimental designs was the absence of confounding alternatives. This is a crucial feature of our design since it allows us to observe violations of structural axioms such as monotonicity or impatience and to reduce the learning effect. The construction of the questions is based on the literature regarding the order effect. A survey can be found in Day et al. (1987). They identify four factors as sources of order effects:

1. **Discovered preference hypothesis** (Plott, 1993) and **Institutional learning**. The first refers to respondents that, when faced with new decisions in unfamiliar environments, exhibit significant randomness in initial decisions. The second is related to the fact that respondents may have never experienced a lab experiment and surely they have never experienced the present design;

2. **Fatigue**: respondents may get tired of the choice task, especially if it is repeated many times. Hence they could exhibit higher randomness in later tasks;

3. **Starting point effect**: Respondents may create a reference point in some choice task and based subsequent choices on that.

Table B.3. **List of Questions - Time**

QUESTIONS		ALTERNATIVES											
1		OS	D										
2		OS	K										
3		OS	I										
4		D	K										
5		D	I										
6	MAIN	K	I										
7		OS	D	K									
8		OS	D	I									
9		OS	K	I									
10		D	K	I									
11		OS	D	K	I								
12		D	I	M(D)									
13	AD	D	I	M(I)									
14		D	I	Im(D)									
15		D	I	Im(I)									
16		Im(OS)	Im(D)	Im(K)	Im(I)								
17		M(OS)	M(D)	M(K)	M(I)								
18		Neu1	Neu2										
19		M(OS)	M(D)	Neu1	Neu2								
20		Im(K)	Im(I)	Neu1	Neu2								
21		D	I	M(D)	M(I)	Im(D)	Im(I)	Neu1	Neu2				
22		OS	D	K	I	M(OS)	M(D)	M(K)	M(I)	Neu1	Neu2		
23	BIG	OS	D	K	I	Im(OS)	Im(D)	Im(K)	Im(I)	Neu1	Neu2		
24		D	I	M(OS)	M(K)	Im(OS)	Im(K)	Neu1	Neu2				
25		OS	D	K	I	M(OS)	M(D)	M(K)	M(I)	Im(OS)	Im(D)	Im(K)	Im(I)

Table B.4. **List of Questions - Risk**

QUESTIONS		ALTERNATIVES									
1		D	S								
2		D	50								
3		D	R								
4		S	50								
5		S	R								
6	MAIN	50	R								
7		D	S	50							
8		D	S	R							
9		D	50	R							
10		S	50	R							
11		D	S	50	R						
12		S	R	F1(S)							
13	AD	S	R	So(S)							
14		S	R	F1(R)							
15		S	R	So(R)							
16		F2(D)	F2(S)	So(50)							
17		F2(D)	F2(R)	So(50)							
18		F2(S)	F2(50)	F2(R)							
19		F2(50)	Sim								
20		D	S	50	R	So(D)	So(S)	So(50)	So(R)	F1(D)	F1(50)
21		D	S	50	R	So(D)	So(S)	So(50)	So(R)		
22	BIG	D	S	50	R	F2(D)	F2(S)	F2(50)	F2(R)		
23		Sim	S	50	R	So(D)	So(S)	So(50)	So(R)		
24		Sim	R	S	50	D	So(D)	So(S)	So(R)		
25		D	S	50	R	F1(D)	F1(S)	F1(50)	F1(R)	Sim	F2(R)

In order to reduce the magnitude of the first two effects, we construct confounding alternatives, as described in Appendix B.1. We also avoid that subjects create an immediate knowledge of the MAIN alternatives designing a heterogenous list of questions. This design feature has been previously exploited by Badescu & Weiss (1987) and suggested by Charness et al. (2012). As described in Table B.3 and B.4, the questions are divided into three main domains: MAIN, AD and BIG. The remaining questions are neutral questions.

As suggested in Carlsson et al. (2012) and Ladenburg & Olsen (2008), we include three trial questions before each part of the experiment in order to reduce institutional learning. These questions have irrelevant alternatives and different numbers of alternatives as to not affect the preferences of individuals or create reference points.

Fatigue is not a major issue. Numerous studies find no or small differences in preferences due to fatigue in experiments with a sequence of identical choice problems (Carlsson et al., 2012) (note that our choice problems are not identical). We set the number of questions to 50, more or less in line with the literature: Agranov & Ortoleva (2017) asked 70 questions, Cavagnaro & Davis-Stober (2014) 120 questions, Manzini et al. (2010) 22 questions, Barberá & Neme (2017) 16 questions. However, the first two experiments involved pairs of lotteries and so they have a simpler design compared to our experiment; while the second two experiments lasted only 15 to 30 minutes which is much less than usual experiments in the field.

Table B.5. **List of Orders**

TIME		RISK	
ORDER 1	ORDER 2	ORDER 1	ORDER 2
21	24	23	21
14	14	13	13
1	4	1	4
23	16	20	24
10	3	10	3
13	13	14	14
3	1	3	1
19	17	16	18
7	8	7	8
22	22	22	22
2	10	2	10
18	18	19	19
11	11	11	11
20	19	17	16
15	15	15	15
9	9	9	9
25	25	25	25
8	6	8	6
17	23	18	20
4	2	4	2
12	12	12	12
16	7	24	7
6	20	6	17
5	5	5	5
24	21	21	23

In Table B.5 we present the four orders that have been used. The construction of the two orders, both in Time and Risk, has followed a structural randomization. Namely, we divided the alternatives in the four similarity groups shown in Table B.3 and B.4. Then, we randomized the positions in order to avoid, as much as possible, similar questions to arise consecutively. In order to test for reference points and, more generally order effects, we fixed the position of certain questions (e.g. number 11) while for others we inverted positions (e.g. number 21 and 24).

In Table B.6 we present the descriptive statistics related to the four possible treatments for WARP violations in the MAIN sets:

1. Order 1 & Time/Risk;

2. Order 1 & Risk/Time;

3. Order 2 & Time/Risk;

4. Order 2 & Risk/Time.

Table B.6. **WARP Main Sets - Treatments**

	TIME		RISK	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Treatment 1	1.8333	2.7826	5.0555	3.8689
Treatment 2	2.7179	3.6343	5.5128	3.4859
Treatment 3	1.5333	2.1613	3.7667	2.9088
Treatment 4	1.6500	2.6365	5.3250	3.1896

NOTES -- In Time no difference in the mean among Treatments is statistically significant. In Risk the 3rd Treatment has less violations than the 1st and 4th with significance at 5%.

B.3 Questionnaire

In order to have more insights into the understanding of the experiment, we conducted a non-incentivized questionnaire. In the first part, summarized in Table B.7, subjects were asked to agree or disagree with some statements. We find that overall subjects show both a good understanding of the experimental design and of the instructions. They also report signals of learning effect expressed both in terms of preference learning and quicker response times. The reward has been considered overall a significant contribution of their daily budget.

Table B.7. **Questionnaire 1**

	QUESTIONS	MEAN	SD
1	**"I got a good understanding of overall experiment (incentives, how to choose, nature of the alternatives)".**	1.552	0.600
2	**"The instructions and explanations provided have been enough to understand experiment's duties".**	1.366	0.622
3	**"The clarity of my preferences on the alternatives improved during the course of the experiment".**	1.993	0.961
4	**"The time required for choosing has reduced during the course of the experiment".**	1.931	1.045
5	**"The money I earn through participating in this experiment is a substantial contribution to my daily budget"**	2.462	1.167

NOTE - Subjects could reply to this statement choosing one of the following: "Strongly Agree", "Agree", "Neutral", "Disagree", "Strongly Disagree". The answers have been matched to numbers from 1 to 5; where 1 is "Strongly Agree" and 5 is "Strongly Disagree".

In the second part of the questionnaire (Table B.8), subjects were asked to reply with a "Yes" or "No" to a series of questions about the structure of the experiment and their behaviour. Overall they accepted the idea that some questions were more difficult than others and they indicated the high number of alternatives as to the main source of complexity. Strangely, but comprehensively, most of the subjects had the impression that some questions were repeated even though it was not the case. They overall confirm to have read all alternatives before choosing; importantly note that reading and paying attention sometimes do not overlap. Finally, they confirm the presence of reference point effects.

Questions 6 and 7 focus on the reasoning of subjects. About half of the subjects report risk neutrality via the use of the "highest expected value" criterion; while about two thirds report patience via the use of the "highest summation" criterion.

Table B.8. **Questionnaire 2**

	QUESTIONS	YES	NO
1	**"Did you find that some questions were more difficult than others?"**	123	22
2	**"If yes, what features did play a main role in making these questions more difficult?" [YES - High number of alternatives] - [NO - Complexity of some alternatives]**	92	53
3	**"Do you think some questions were asked multiple times?"**	133	12
4	**"Did you always read all the alternatives carefully before choosing?"**	118	27
5	**"Did you base some decisions on previous choices?"**	123	22
6	**"In case of lotteries, in general, did you calculate the expected value and choose according to the "highest expected value" criterion?"**	75	70
7	**In case of delayed payment plans, in general, did you calculate the summation of the plan and choose according to the "highest summation" criterion?**	97	48

NOTE - Question 2 wasn't given to subjects as a YES/NO; but here it has been translated into YES/NO to ease the Table representation.

B.4 Instructions

Figure B.1. **General Instructions**

General Instructions

This is an experiment in the economics of decision-making and in particular risk and time attitudes of individuals.

Notice there is no RIGHT or WRONG answer for any of these questions. We are interested in studying your preferences.

The tasks are extremely simple and if you make good decisions you may earn a **considerable amount of money** that will be **paid to you by direct debit** at the end of the experiment and at particular times in the future.

The currency in this experiment is called **tokens.**

The experiment should last at maximum 1:30 hour even though more probably around 1 hour. It is divided in four parts:

1. **Part 1:** 25 Questions;
2. **Part 2:** 25 Questions;
3. **Questionnaire**
4. **Test**

At the end of experiment, the computer will randomly choose on question for each part of the experiment and pay your choice. In case of Lottery the computer will also play out the Lottery according to the respective probabilities.

Your total earning from the experiment will consist of the sum of three components:

1. One Choice from Part 1;
2. One Choice from Part 2;
3. A participation fee of £ 5.

Notice:

- There is no time constraint for any question;
- Before each part of the experiment you will complete a three questions trial in order to make you fully understand which kind of questions you are going to answer. The alternatives in the trial questions do not have sense and will not be considered in the calculation of your reward.
- After you click the "OK" button, there might be a short delay before the next question appears, due to the software. Please be patient.
- At the end of **PART 1** the screen **"BREAK"** will appear. Please remain sit and wait. During the break you are not allowed to speak with anyone. You will receive the instructions for **PART 2** and the experiment will restart.

Figure B.2. **Specific Instructions I**

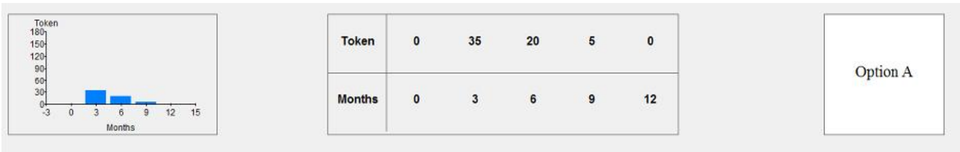
Specific Instructions

Instructions are visualized on the screen before each Part of the Experiment. This is the paper version. You can keep it and revise it in case you have doubts about the experiment.

The exchange rate for this part of the experiment is:

20 Token = £ 1

In this part, you will choose among Delayed Payment Plans. The following screen shows a delayed payment plan:

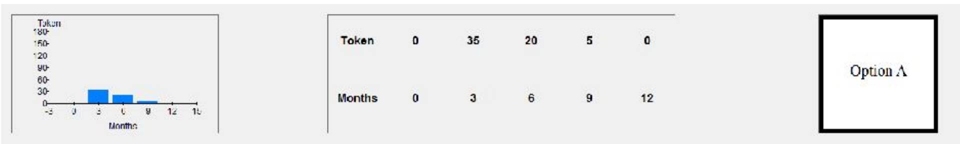


A delayed payment plan is described in two ways:

- A histogram, on the left, allows to visualize the number of tokens paid at different months;
- A table in which Token and Months are written.

This delayed payment plan allows you to win 35 Tokens in 3 months, 20 Tokens in 6 months and 5 Tokens in 9 months.

On the right you can see a Button to select. If you click on “Option A” you select this Delayed Payment Plan and a black frame will appear.



You can change your choice even after you have selected one. When you are sure of your choice you can click the Button “OK” in the Right-Bottom part of the screen (see below). Once you clicked “OK” your decision is recorded and you cannot change it.

Notice:

- Questions may have a different number of alternatives to choose from;
- Delayed Payment Plans may different both in Token and in Time of Payments.

Figure B.3. **Specific Instructions II**

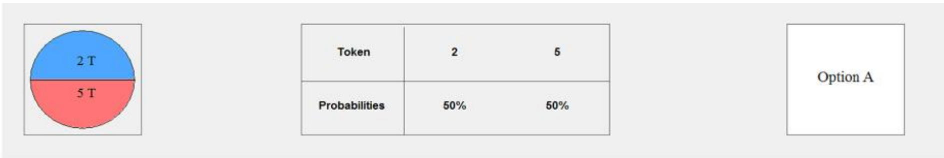
Specific Instructions

Instructions are visualized on the screen before each Part of the Experiment. This is the paper version. You can keep it and revise it in case you have doubts about the experiment.

The exchange rate for this part of the experiment is:

10 Token = £ 1

In this part, you will choose among Lotteries. The following screen shows a lottery:

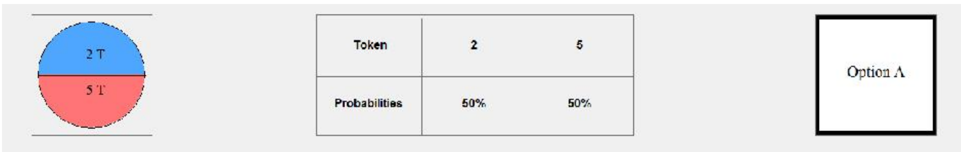


A lottery is described in two ways:

- A pie, on the left, allows to visualize the probabilities to win;
- A table in which Token and Probabilities are written.

This lottery allows you to win 2 Token with 50% probability and 5 Token with 50% probability.

On the right you can see a Button to select. If you click on “Option A” you select this Lottery and a black frame will appear.



You can change your choice even after you have selected one. When you are sure of your choice you can click the Button “OK” in the Right-Bottom part of the screen (see below). Once you clicked “OK” your decision is recorded and you cannot change it.

Notice:

- Questions may have a different number of alternatives to choose from;
- Lotteries may different both in Token and in Probabilities.

B.5 Screens

Figure B.4. **Screen Lotteries**

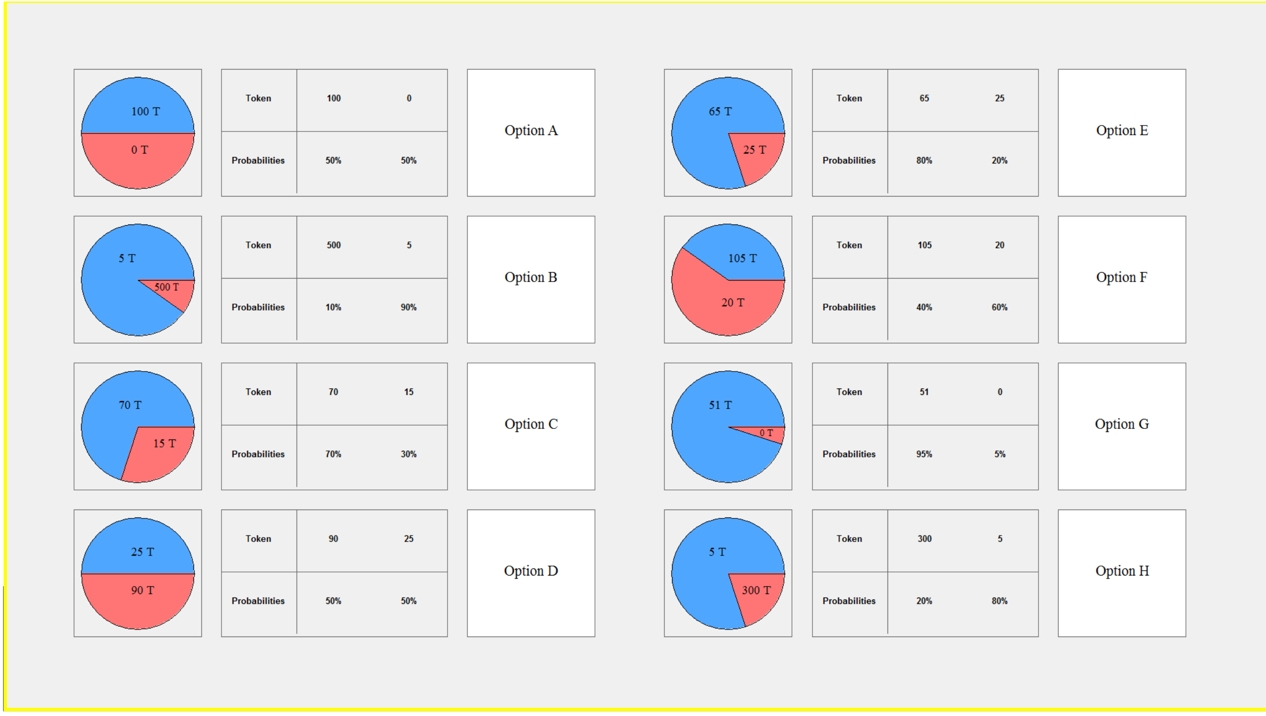


Figure B.5. **Screen Delayed Payment Plans**

Appendix C

Appendix to Chapter 3

C.1 Additional results

C.1.1 More on Higher, Lower Negative Transitivity and Strong Stochastic Transitivity.

This section helps to clarify the two new deterministic versions of transitivity that we introduced: LNT and HNT. So far we proved that LNT neither implies nor is implied by transitivity; while HNT is stronger than transitivity. Interestingly, a binary relation that satisfies this property is a special case of semiorder. In fact, it satisfies both strong intervality and semitransitivity.

Definition. *A binary relation $\succ$ satisfies strong intervality [SI] if $x \succ y$ and $z \succ w$ implies $x \succ w$ or $z \succ y$ for all x, y, z, w.*

Definition. *A binary relation $\succ$ satisfies semitransivity [ST] if $x \succ y$ and $y \succ z$ implies $x \succ w$ or $w \succ z$ for all x, y, z, w.*

Proposition. *If $\succ$ satisfies HNT then it satisfies SI.*

Proof. Suppose by contradiction that $x \succ y$, $z \succ w$, $x \not\succ w$ and $z \not\succ y$. If $w = y$, the contradiction arise from $x \succ y$ and $x \not\succ y$. If $w = z$, HNT is violated since $x \succ y$, $x \not\succ z$ and $z \not\succ y$. If $w = x$, transitivity is violated since $z \succ x \succ y$ but $z \not\succ y$.

Conversely, let $x \succ y$, $z \succ w$ and $x \succ w$. SI is satisfied but not HNT. $\square$

Proposition. *If $\succ$ satisfies HNT then it satisfies ST.*

Proof. Suppose by contradiction that $x \succ y$, $y \succ z$, $x \not\succ w$ and $w \not\succ z$. If either $w = x$ or $w = z$ then transitivity is violated since $x \not\succ z$ and if $w = y$ again the contradiction arise from $x \succ y$ and $x \not\succ y$.

Conversely, let $x \succ y$ then ST is satisfied but not HNT. □

Note that, HNT and LNT are not exactly properties of a binary relation. They are properties of a tuple $(\succ, \gg)$, with $\gg$ that preserves $\succ$. Formally, $x \succ y$ implies $x \gg y$ for all x, y. Fishburn (1999) shows that a function satisfying this latter property can be constructed on any set endowed with an acyclic binary relation $\succ$. It is therefore not surprising that $x \succ y \succ z$, $x \not\succ z$ satisfies LNT but not transitivity and that a binary relation satisfies either LNT or HNT if and only if it is acyclic.

One can notice that HNT characterizes a binary relation that is stronger than both interval order and semiorder. This condition is respected in stochastic choice. The reader may verify in Fishburn (1973) that SST is stronger of both Interval Stochastic Transitivity (IST) and J-Stochastic Transitivity (JST) that together characterize $\succ_\mu$ to be a semiorder for all $\mu \in [0.5, 1)$, while only IST characterizes $\succ_\mu$ to be an interval order for all $\mu \in [0.5, 1)$. The definition of $\succ_\mu$ can be found in Section 3.4.2 of Chapter 3.

C.1.2 Relaxing Antisymmetry

In Section 3.3 we introduced two properties MST* and SST* that are modifications of the well-known MST and SST. Here, we generalize this modification to WST that is traditionally defined as follows:

Definition. *A stochastic choice rule p satisfies Weak Stochastic Transitivity [WST] if for all $x, y, z \in A$:*

$$p(x,y) \ \ \& \ \ p(y,z) \geq 0.5 \quad \Rightarrow \quad p(x,z) \geq 0.5$$

This property is equivalent to the existence of a utility function u such that $u(x) > u(y)$ if and only if $p(x,y) > 0.5$ and $u(x) = u(y)$ if and only if $p(x,y) = 0.5$. Marschak & Block (1960) presents a comprehensive analysis of WST. We now define a modification of the property.

Definition. *A stochastic choice rule p satisfies Weak Stochastic Transitivity* [WST*] if for all* $x,y,z \in A$*:*

$$p(x,y) \ \& \ p(y,z) \ \& \ p(x,z) \geq \frac{1}{2}$$

and either of the following holds:

(1) $\min[p(x,y),p(y,z)] = \frac{1}{2} \quad \Rightarrow \quad p(x,z) > \min[p(x,y),p(y,z)]$

(2) $p(x,z) = p(x,y) = p(y,z)$

The following result states that if AST is satisfied then the two above properties are equivalent.

Proposition. *A stochastic choice rule p satisfies WST if and only if it satisfies WST*.*

Proof. Let AST be satisfied.. The following cases are possible:

(i) $p(x,y) = p(y,z) = p(x,z) = 0.5$. Both WST and WST* are not violated;

(ii) $p(x,y) \ \& \ p(x,z) > 0.5$ and $p(y,z) = 0.5$. WST and WST* are both not violated;

(iii) $p(x,y) > p(y,z) = p(x,z) = 0.5$. Both WST and WST* are violated since $u(x) = u(z) = u(y) < u(x)$ and condition (1) of WST* is violated. Note that AST is not violated.

(iv) $p(x,y) \ \& \ p(y,z) > 0.5$ and $p(x,z) = 0.5$. Both WST and WST* are violated. Again, AST is satisfied.

(v) $p(x,y) \ \& \ p(y,z) \ \& \ p(x,z) > 0.5$. Both WST and WST* are satisfied.

□

C.2 Proofs

C.2.1 Proof of Theorem 1

Necessity: Suppose ε satisfies triangle inequality and take $x,y,z \in A$. Let $x \succ y$ and $y \succ z$ then:

$$u(x) - u(y) > \varepsilon(x,y)$$

$$u(y) - u(z) > \varepsilon(y,z)$$

Combining the two inequalities and using triangle inequality:

$$u(x) - u(z) > \varepsilon(x,y) + \varepsilon(y,z)$$

$$u(x) - u(z) > \varepsilon(x,z)$$

hence $x \succ z$ and transitivity is satisfied.

Sufficiency: The proof is constructive and it is divided into five steps.

Step 3.

We first construct a utility function such that $x \succ y$ implies $u(x) > u(y)$. Let $u(x) = |\{t \in A : t \not\succ x\}|$.

By $x \succ y$ and transitivity of $\succ$ we have $\{t \in A : t \succ x\} \subsetneq \{t \in A : t \succ y\}$ since if an element t is preferred to x, by transitivity it is also preferred to y. Hence, the completements of these sets have an opposite subset relation: $\{t \in A : t \not\succ x\} \supsetneq \{t \in A : t \not\succ y\}$, which implies $u(x) > u(y)$.[1].

Step 4.

From this step on, we construct the quasi-metric. First, let $d(x,y)$ be the shortest path metric of the transitive graph G of the partial order $\succ$. This metric has the following characteristics:

1. $x \succ y \Rightarrow d(x,y) = 1$
2. $[x \not\succ y \wedge y \not\succ x] \Rightarrow d(x,y) > 1$
3. $d(x,y) > 2$ implies that there exists a sequence $(x_1, \ldots x_n)$ such that:

$$x_1 \succ x_2 \prec x_3 \succ x_4 \cdots \succ x_n \prec x_{n+1}$$

[1]This step has been highlighted also by Fishburn (1999)

where n is even.

Point (1), (2) are immediate. Let's prove point (3):

Suppose $d(x,y)=2$ then there exists an element z such that either (i) $x \succ z, y \succ z$ or (ii) $z \succ x, z \succ y$. Clearly, if $x \succ z \succ y$ then by transitivity $x \succ y$ but then $d(x,y)=1$.

(i) - Suppose we want to increase $d(x,y)$ then take a w: if $w \succ x$ (or $w \succ y$) by transitivity $w \succ z$ and $d(x,y)=2$. If $z \succ w$ then $d(x,y)$ is not influenced.

(ii) - Again take a w: if $w \succ z$ then by transitivity $w \succ x, w \succ y$ and $d(x,y)=2$. If $z \succ w$ then $d(x,y)$ is not influenced.

Hence, we need two elements z,w such that $x \succ z$, $w \succ z$ and $w \succ y$ given $d(x,y)=3$. Following this idea we can increase $d(x,y)=n-1$ where n is the cardinality of A.

Step 5.

The shortest path metric proposed in Step 2 needs to be refined. We transform the graph G into an undirected weighted graph. For any edge $e=(x,y)$ of the undirected graph corresponding to the directed graph G define a weight $w(e)=|u(x)-u(y)|$. Define $P(x,y)$ as a path from x to y and define a weight $w[P(x,y)]=\sum_{e\in P(x,y)} w(e)$. Let $\delta(x,y)$ be the minimum weighted path from x to y. Since $\succ$ is not complete, it can be that there is no path between two elements x,y. If this is the case we set $\delta(x,y)=n^*$ where $n^*=\max_{x,y}\delta(x,y)+\varepsilon$ with $\varepsilon>0$. This choice respects the finiteness of δ and it is in line with the usual convention of setting the distance between two unconnected node as infinite. By Monjardet (1980) we have that $\delta(x,y)$ is a quasi-metric. This result is straightforward:

(i) $\delta(x,x)=0$;

(ii) $\delta(x,y)=\delta(y,x)$ since the graph is undirected;

(iii) $\delta(x,z)\leq\delta(x,y)+\delta(y,z)$. To see this suppose y is on the Minimum Weighted Path between x,z then by definition $\delta(x,y)+\delta(y,z)=\delta(x,z)$. If y is not on the Minimum Weighted Path then by definition $\delta(x,z)<\delta(x,y)+\delta(y,z)$ otherwise minimality would be violated. Finally, if x,z are not connected there exists no y such that both x,y are connected and y,z are connected. This means that if x,y are not connected $\delta(x,y) = \delta(x,z)$ and by positivity $\delta(y,z) \geq 0$, hence triangle inequality is satisfied.

Step 6.

Now, we need to prove that $\succ$ has a BT representation. First, we prove that $x \succ y$ implies $u(x) - u(y) > \gamma \cdot \delta(x,y)$ for all $\gamma \in [0,1)$. This is immediate. By Step 1 we have $u(x) > u(y)$. By definition, $\delta(x,y) = u(x) - u(y)$. Hence, $u(x) - u(y) > \gamma \cdot |u(x) - u(y)|$ for any $\gamma \in [0,1)$.

Step 7.

Finally, we have to prove that $[x \not\succ y \wedge y \not\succ x]$ implies $u(x) - u(y) \leq \gamma \cdot \delta(x,y)$ for some $\gamma \in (0,1]$.

Suppose w.l.o.g. that $u(x) \geq u(y)$ and let's prove by induction over the shortest path $d(x,y)$ as defined in Step 2 from x to y.

By Step 2, we start analysing the case $d(x,y) = 2$. We have two cases: (i) $x \succ z$, $y \succ z$ and (ii) $z \succ x$, $z \succ y$.

(i) - if $u(x) = u(y)$ then we can write the inequality as $u(x) - u(z) + u(y) - u(z) \geq 0$. Since $u(x) > u(z)$ and $u(y) > u(z)$ the inequality is strict and for all $\gamma \in [0,1)$ the inequality $\gamma \cdot \delta(x,y) \geq 0$ is satisfied.

If $u(x) > u(y)$ then the inequality is $u(x) - u(z) + u(y) - u(z) \geq u(x) - u(y)$ but $2u(y) \geq 2u(z)$ and the inequality is strict. Hence, there exists a $\gamma \in (0,1)$ such that the weak inequality is satisfied. The value of γ is:

$$1 > \gamma \geq \frac{u(x) - u(y)}{u(x) + u(y) - 2u(z)} > 0$$

(ii) - if $u(x) = u(y)$ by the same argument as before, for all $\gamma \in (0,1)$ we have that $\gamma \cdot [2u(z) - u(x) - u(y)] \geq 0$.

If $u(x) > u(y)$ then there exists a $\gamma \in (0,1)$ such that the inequality is satisfied and the value is:

$$1 > \gamma \geq \frac{u(x) - u(y)}{2u(z) - u(x) - u(y)} > 0$$

By definition of $\delta(x,y)$ if there exist two elements z,t such that conditions (i) and (ii) are satisfied then we select the minimum path. Furthermore, again by definition of δ, $\delta(x,y)$ is minimized and so γ is maximized for each x,y. However, this is not enough, in fact in order to satisfy triangle inequality for all $x,y \in A$ we have to multiply $\delta(x,y)$ by the same parameter γ. The selection is found maximizing among alternatives:

$$\gamma^* = \underset{x,y \in A:[x \not\succ y \wedge y \not\succ x]}{\operatorname{argmax}} \frac{u(x) - u(y)}{\delta(x,y)}$$

This condition guarantees that if γ^* is selected by x,y then $[x \not\succ y \wedge y \not\succ x]$ implies $[r \not\succ t \wedge t \not\succ r]$ for all other $r,t \in A$.

Suppose the result holds for $d(x,y) = n$ and let $d(x,y) = n+1$. By Step 2, there is a sequence $(x_1, \ldots, x_{n+2})$ such that:

$$x_1 \succ x_2 \prec x_3 \succ x_4 \cdots \succ x_{n+1} \prec x_{n+2}$$

The inequality that must be satisfied is:

$$u(x_1) - u(x_{n+2}) \leq \sum_{i,j} u(x_i) - u(x_j)$$

where i is odd, j is even and i,j are consecutive numbers.

If $n+2$ is even then the inequality is immediately satisfied. If $n+2$ is odd then the inequality is:

$$0 \leq -u(x_2) + u(x_3) - u(x_2) + u(x_3) - u(x_4) + u(x_5) \ldots u(x_{n+2}) - u(x_{n+1})$$

if the condition holds for $d(x,y) = n$ then the inequality is reduced to:

$$0 \leq u(x_{n+2}) - u(x_{n+1}) + u(x_{n+2}) - u(x_{n+1})$$

which clearly holds strictly since $u(x_{n+2}) > u(x_{n+1})$.

Hence, we have proved that the representation holds with $\varepsilon(x,y) = \gamma^* \cdot \delta(x,y)$

for all $x,y \in A$. Clearly, $\varepsilon(x,y)$ maintain all the properties of $\delta(x,y)$ and it is therefore a quasi-metric.

C.2.2 Proof of Theorem 2

Necessity: Let $u(x) > u(y) > u(z)$ and suppose by contradiction that:

$$p(x,z) \leq \min[p(x,y), p(y,z)]$$

and it is not the case that $p(x,y) = p(y,z) = p(x,z)$.

Since c' is strictly increasing in p, given four real numbers $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \Re^+$:

$$c'(p(x,z)) + \lambda_1 = c'(p(x,y))$$

$$c'(p(x,z)) + \mu_1 = c'(p(y,z))$$

$$c'(p(z,x)) - \lambda_2 = c'(p(y,x))$$

$$c'(p(z,x)) - \mu_2 = c'(p(z,y))$$

By definition, the following holds:

$$c'_{\{x,y\}}(p(x,y)) - c'_{\{x,y\}}(p(y,x)) + c'_{\{y,z\}}(p(y,z)) - c'_{\{y,z\}}(p(z,y)) =$$

$$= c'_{\{x,z\}}(p(x,z)) - c'_{\{x,z\}}(p(z,x))$$

and

$$\eta(x,y) \cdot [c'(p(x,y)) - c'(p(y,x))] + \eta(y,z) \cdot [c'(p(y,z)) - c'(p(z,y))] =$$

$$= \eta(x,z) \cdot [c'(p(x,z)) - c'(p(z,x))]$$

Substituting, we obtain:

$$[\eta(x,y) + \eta(y,z)] \cdot [c'(p(x,z)) - c'(p(z,x))] + \eta(x,y) \cdot (\lambda_1 + \lambda_2) + \eta(y,z) \cdot (\mu_1 + \mu_2) =$$

$$= \eta(x,z) \cdot [c'(p(x,z)) - c'(p(z,x))]$$

If the inequality is strict then $\eta(x,z) > \eta(x,y) + \eta(y,z)$ violating triangle inequality. If $p(x,y) = p(x,z) < p(y,z)$ or $p(y,z) = p(x,z) < p(x,y)$ again triangle inequality is violated.

Sufficiency: Let MST* be satisfied. By Theorem 1 of He & Natenzon (2018): there exists a moderate utility representation, namely a utility function u and a distance metric d such that for any $x,y,z,w \in A$ (note that their proof is constructive):

$$p(x,y) \geq p(z,w) \Leftrightarrow \frac{u(x)-u(y)}{d(x,y)} \geq \frac{u(z)-u(w)}{d(z,w)}$$

We construct a strictly increasing function F that maps the probabilities into the numerical representation given by u,d such that:

$$F[p(x,y)] = \frac{u(x)-u(y)}{d(x,y)}$$

$$F[p(z,w)] = \frac{u(z)-u(w)}{d(z,w)}$$

The construction uses a piecewise linear function F on the interval $[0.5,1)$. First let $F(0.5) = 0$. Then, for all $x,y \in A$ define an interval $[p(x,y) - \varepsilon, p(x,y) + \varepsilon]$ with ε sufficiently small such that the interval does not overlap with any other intervals.

Let $F[p(x,y) - \varepsilon] = \frac{u(x)-u(y)}{d(x,y)} - \delta$ and $F[p(x,y) + \varepsilon] = \frac{u(x)-u(y)}{d(x,y)} + \delta$ again with δ sufficiently small such that no interval $\left[\frac{u(x)-u(y)}{d(x,y)} - \delta, \frac{u(x)-u(y)}{d(x,y)} + \delta\right]$ overlaps with other intervals.

Take four elements $x,y,z,w \in A$ such that $p(x,y) > p(z,w)$ and the probabilities are immediate successors. For all $p \in [p(z,w) + \varepsilon, p(x,y) - \varepsilon]$ define the following linear function F passing through the points above defined:

$$\frac{F(p) - F[p(z,w)+\varepsilon]}{F[p(x,y)-\varepsilon] - F[p(z,w)+\varepsilon]} = \frac{p - [p(z,w)+\varepsilon]}{p(x,y) - \varepsilon - [p(z,w)+\varepsilon]}$$

The same function is defined for intervals $[0.5, p(z,w) - \varepsilon]$ and $[p(x,y) + \varepsilon, 1)$ where $p(z,w)$ and $p(x,y)$ are respectively the minimum and maximum probabilities.

Finally, for all $x, y \in A$ in the interval $(p(x,y) - \varepsilon, p(x,y) + \varepsilon)$ define another linear function F:

$$\frac{F(p) - F[p(x,y) - \varepsilon]}{F[p(x,y) + \varepsilon] - F[p(x,y) - \varepsilon]} = \frac{p - [p(x,y) - \varepsilon]}{p(x,y) + \varepsilon - [p(x,y) - \varepsilon]}$$

which can be rewritten as:

$$F(p) = p - p(x,y) + \frac{u(x) - u(y)}{d(x,y)}$$

Clearly, we have $\lim_{p \to p(x,y)} F(p) = \frac{u(x)-u(y)}{d(x,y)}$. By the existence of the limit in $p(x,y)$ for all $x, y \in A$, F is also differentiable in the relevant finite points $p(x,y)$. Also F is strictly increasing on $[0.5, 1)$ and bounded by finiteness of u and strict positivity of d, hence F is integrable.

Let a cost function be $c(p) = \int_{0.5}^{p} F(t)dt$. By the First Fundamental Theorem of Calculus, $c'(p) = F(p)$.

Finally, we can use u, d from He and Natenzon (2018) and c just defined to write a maximization problem:

$$\max_{p(x,y)} u(x)p(x,y) + u(y)[1 - p(x,y)] - d(x,y)c(p(x,y))$$

which can be rewritten (dividing everything by $d(x,y) > 0$) as:

$$\max_{p(x,y)} p(x,y)\frac{[u(x) - u(y)]}{d(x,y)} + \frac{u(x)}{d(x,y)} - c(p(x,y))$$

By the FOCs we have:

$$\frac{u(x) - u(y)}{d(x,y)} = c'(p(x,y))$$

This proves that MST* implies the existence of a BAPU representation where η is a metric.

C.2.3 Proof of Theorem 3

Necessity: Let $u(x) - u(z) > \varepsilon(x,z)$. By the violation of triangle inequality, $u(x) - u(y) + u(y) - u(z) > \varepsilon(x,y) + \varepsilon(y,z)$ which implies either $u(x) - u(y) > \varepsilon(x,y)$ or

$u(y) - u(z) > \varepsilon(y,z)$.

Sufficiency: Using Theorem 2.5 and Theorem 4.1 by Aleskerov et al. (2007) the following results hold: the existence of a BT representation if and only if $\succ$ is acyclic if and only if there exists a weak order $\gg$ such that $x \succ y$ implies $x \gg y$. Namely, $\gg$ preserves $\succ$. Therefore, $(A, \gg)$ is a weakly ordered set. We can define LNT on $(A, \gg)$. We now need to construct a function ε that violates triangle inequality (or satisfy it with equality) wherever LNT is satisfied at some $x,y,z \in A$.

If $x \succ y$ let $\varepsilon(x,y) = |u(x) - u(y)| - \sigma$ for some small $\sigma > 0$. If $[x \not\succ y \wedge y \not\succ x]$ let $\varepsilon(x,y) = |u(x) - u(y)|$. The binary relation is represented:

1. $x \succ y$ implies $u(x) - u(y) > \varepsilon(x,y)$;

2. $[x \not\succ y \wedge y \not\succ x]$ implies $u(x) - u(y) = \varepsilon(x,y)$.

We have to prove that ε violates triangle inequality when $u(x) \geq u(y) \geq u(z)$. Note that $|u(x) - u(y)| + |u(y) - u(z)| = |u(x) - u(z)|$. Also, for all $x,y \in A$, $u(x) = u(y)$ implies $x \not\succ y$ and vice versa; but the converse is not true. For instance, let $x \succ y$, $y \succ z$ but $x \not\succ z$. One can see that $x \gg y \gg z$ preserves $\succ$, hence $u(x) > u(z)$.

By LNT the following cases are possible:

(i) if $x \succ z$, $x \succ y$ and $y \succ z$ then: $|u(x) - u(z)| - \sigma = \varepsilon(x,z) > \varepsilon(x,y) + \varepsilon(y,z) = |u(x) - u(z)| - 2\sigma$;

(ii) if $x \succ z$ and $x \succ y$ then: $|u(x) - u(z)| - \sigma = \varepsilon(x,z) = \varepsilon(x,y) + \varepsilon(y,z) = |u(x) - u(z)| - \sigma$;

(iii) if $x \succ y$ and $y \succ z$ then: $|u(x) - u(z)| = \varepsilon(x,z) > \varepsilon(x,y) + \varepsilon(y,z) = |u(x) - u(z)| - 2\sigma$;

(iv) if $x \succ y$ then: $|u(x) - u(z)| = \varepsilon(x,z) > \varepsilon(x,y) + \varepsilon(y,z) = |u(x) - u(z)| - \sigma$;

(v) if $\succ = \emptyset$ then $\varepsilon(x,z) = \varepsilon(x,y) + \varepsilon(y,z)$.

C.2.4 Proof of Theorem 4

Necessity Let $u(x) > u(y) > u(z)$ and suppose by contradiction that:

$$p(x,z) \geq \max[p(x,y), p(y,z)]$$

Since c' is increasing in p, given four real numbers $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \Re^+$:

$$c'(p(x,z)) - \lambda_1 = c'(p(x,y))$$

$$c'(p(x,z)) - \mu_1 = c'(p(y,z))$$

$$c'(p(z,x)) + \lambda_2 = c'(p(y,x))$$

$$c'(p(z,x)) + \mu_2 = c'(p(z,y))$$

Following the same approach that we used in Theorem 2 we obtain:

$$[\eta(x,y) + \eta(y,z)] \cdot [c'(p(x,z)) - c'(p(z,x))] - \eta(x,y) \cdot (\lambda_1 + \lambda_2) - \eta(y,z) \cdot (\mu_1 + \mu_2) =$$

$$= \eta(x,z) \cdot [c'(p(x,z)) - c'(p(z,x))]$$

hence: $\eta(x,z) < \eta(x,y) + \eta(y,z)$ satisfies triangle inequality.

Note that if either $p(x,z) = p(x,y) > p(y,z)$ or $p(x,z) = p(y,z) > p(x,y)$. Then we have $\lambda_1 = \lambda_2 = 0$ but $\mu_1, \mu_2 > 0$ (or vice versa) and the result still holds.

Sufficiency This part requires a construction similar to the one in He & Natenzon (2018). The utility function is taken from their construction: WST is satisfied; hence we let $x \succ y$ if and only if $p(x,y) > 0.5$, obtaining a linear order. Since A is finite, there exists a utility function $u : A \to \{1, 2, \ldots, |X|\}$ such that $x \succ y$ if and only if $u(x) > u(y)$.

Let $Y = \mathcal{A}_2$ be the set of binary subsets of A and let m be the cardinality of the set $\{|p(x,y) - 0.5| > 0 : \{x,y\} \in Y\}$. Partition the set Y in m disjoint sets such that $\{x,y\} \in Y_i$ and $\{z,w\} \in Y_j$ with $i \geq j$ if $p(z,w) \geq p(x,y) > 0.5$. Let $(D_i)_{i=1}^m$ be a sequence of strictly positive numbers that is attached to the sequence of sets $(Y_i)_{i=1}^m$. The distance between x, y when $\{x,y\} \in Y_i$ is defined as follows:

$$\eta(x,y) = D_i|u(x) - u(y)|$$

Take three elements x,y,z such that $u(x) > u(y) > u(z)$, with $\{x,y\} \in Y_i$, $\{y,z\} \in Y_j$ and $\{x,z\} \in Y_l$, then:

$$\eta(x,y) + \eta(y,z) - \eta(x,z) =$$

$$= D_i|u(x) - u(y)| + D_j|u(y) - u(z)| - D_l|u(x) - u(y) + u(y) - u(z)| =$$

$$= (D_i - D_l)|u(x) - u(y)| + (D_j - D_l)|u(y) - u(z)|$$

We will construct the sequence $(D_i)_{i=1}^m$ recursively. Note that we want triangle inequality to be (weakly) violated:

$$(D_i - D_l)|u(x) - u(y)| + (D_j - D_l)|u(y) - u(z)| \leq 0$$

$$(D_l - D_i)|u(x) - u(y)| \geq (D_j - D_l)|u(y) - u(z)|$$

Given the above inequality we need to identify the worst case scenario. Namely, the case that, when satisfied, guarantees that the inequality is satisfied in all other cases. We argue that the worst case scenario is when utilities are $u(x) = n$, $u(y) = n - 1$, $u(z) = 1$ (where $n = |A|$). Probabilities are such that $p(y,z)$ is minimum, therefore D_j is maximum; $p(x,z) < p(x,y)$ and these latter are immediate successor.

To verify that this is the worst case scenario note the following two facts:

- The LHS is strictly increasing in $D_l - D_i$; hence for them being immediate successor is the worst case. In fact, the LHS is increasing in D_l, while RHS is decreasing in D_l, hence when this term is minimum, $D_l = f(l+1)$, it is the worst case;

- The LHS is decreasing in $u(y)$, while the RHS is increasing in $u(y)$ therefore when $u(y)$ is maximum, $u(y) = n - 1$, it is again the worst case.

Note also that m is bounded above by $\binom{n}{2}$ which is the number of binary subsets of a set with cardinality n. We can thus rewrite the above inequality in order to find the function f that maps the sequence $\{1,\dots i,\dots,m\}$ into positive real numbers giving us the desired sequence $(D_i)_{i=1}^m$:

$$[f(l+1)-f(l)]|u(x)-u(y)| \geq \left[f\left(\binom{n}{2}\right)-f(l+1)\right]|u(y)-u(z)|$$

$$[f(l+1)-f(l)] \geq \left[f\left(\binom{n}{2}\right)-f(l+1)\right](n-2)$$

We assume that this inequality is satisfied with equality at the worst case scenario. Therefore it becomes a simple difference equation. To solve it, we need to define the initial conditions; or in other words the first step. For instance, let $f(l)=1$ and $f(l+1)=2$ for $l=1$. Solving the worst case scenario with equality becomes:

$$1=\left[f\left(\binom{n}{2}\right)-2\right](n-2)$$

$$f\left(\binom{n}{2}\right)=\frac{1}{n-2}+2$$

$f(l)$ can be recursively constructed for any $l \geq 3$. Solving the equation:

$$f(l+1)-f(l)=\left[\frac{1}{n-2}+2-f(l+1)\right](n-2)$$

$$f(l+1)=\frac{1+2(n-2)+f(l)}{1+(n-2)}$$

Now we verify that triangle inequality is always violated when Strong Stochastic Transitivity* is violated:

(i) If either $p(y,z)>p(x,z)=p(x,y)$ or $p(x,y)>p(x,z)=p(y,z)$; $D_l=D_i>D_j$ or $D_l=D_j>D_i$ and $\eta(x,y)+\eta(y,z)-\eta(x,z)<0$;

(ii) If either $p(x,y)>p(y,z)>p(x,z)$ or $p(y,z)>p(x,y)>p(x,z)$; $D_l>D_j>D_i$ or $D_l>D_i>D_j$ and $\eta(x,y)+\eta(y,z)-\eta(x,z)<0$;

(iii) If either $p(y,z) > p(x,z) > p(x,z)$ or $p(x,y) > p(x,z) > p(y,z)$ then again $\eta(x,y) + \eta(y,z) - \eta(x,z) \leq 0$, where the equality is satisfied in the worst case scenarios.

If triangle inequality is violated then it is satisfied in the two opposite directions, namely $\eta(z,x) + \eta(x,y) - \eta(z,y) \geq 0$ and $\eta(y,z) + \eta(z,x) - \eta(y,x) \geq 0$. We need to verify that these conditions are met when SST* is satisfied:

$$\eta(z,x) + \eta(x,y) - \eta(y,z) = (D_i + D_l)|u(x) - u(y)| + (D_l - D_j)|u(y) - u(z)| > 0$$

$$\eta(y,z) + \eta(z,x) - \eta(y,x) = (D_l - D_i)|u(x) - u(y)| + (D_l + D_j)|u(y) - u(z)| > 0$$

since $D_l \geq D_j$ and $D_l \geq D_i$.

Now, let $p(x,y) > p(z,w)$ with $\{x,y\} \in Y_i$ and $\{z,w\} \in Y_j$:

$$\frac{u(x) - u(y)}{\eta(x,y)} = \frac{1}{D_i} > \frac{1}{D_j} = \frac{u(z) - u(w)}{\eta(z,w)}$$

Let $p(x,y) = p(z,w)$ with $\{x,y\}, \{z,w\} \in Y_i$:

$$\frac{u(x) - u(y)}{\eta(x,y)} = \frac{1}{D_i} = \frac{1}{D_i} = \frac{u(z) - u(w)}{\eta(z,w)}$$

To complete the proof a cost function can be constructed as in Theorem 2 in order to provide the BAPU representation.

C.3 The special case of 3 elements

Theorem.

(i) Let $|A| = 3$; the result in Theorem 2 holds for every strictly convex, C^1 cost function;

(ii) If $|A| > 3$, the result at (i) does not hold.

Proof. - (i)

Let c be **any** strictly convex, C^1 function with $\lim_{p\to 0} c(p) = -\infty$. We have that c' is strictly increasing in p. Therefore $c'(p(x,y)) \geq c'(p(y,x))$ if and only if $p(x,y) \geq p(y,x)$.

We need to prove that there exists a u and η such that this latter is a metric and p is the outcome of a perturbed maximization problem. This proposition is equivalent to the existence of solutions for a system inequalities.

In order to prove the compatibility of such system we can rely on Motkin theorem of alternatives; Dantzig (1963):

Theorem (Motzkin's Theorem)**.** *Let A, B, C be given matrices with A being non vacuous. Then one and only one of the following is feasible.*

- *There exists x such that $Ax > 0$, $Bx \geqq 0$, $Cx = 0$.*
- *There exists π, μ, γ such that $A^T\pi + B^T\mu + C^T\gamma = 0$, $\pi \geq 0$ and $\mu \geqq 0$.*

where $\geq$ means semi-positive and $\geqq$ means non-negative.

By Weak Stochastic Transitivity, and and assuming $p(x,y) \neq 0.5$ for all x,y, we can define a linear order over the elements of X such that $u(x) > u(y)$ if and only if $p(x,y) > 0.5$. Let $u = (1,2,3\ldots,n)$ where $u(x) = 1$ if x is the worst element and $u(x) = n$ if it is the best element. So, take w.l.o.g. three consecutive elements $x,y,z \in A$ such that $u(x) > u(y) > u(z)$ and let Moderate Stochastic Transitivity be satisfied. For simplicity, let $c^{'}(p(x,y)) - c^{'}(p(y,x)) = a$, $c^{'}(p(y,z)) - c^{'}(p(z,y)) = c$ and $c^{'}(p(x,z)) - c^{'}(p(z,x)) = b$. Then the system is as follows:

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta(x,y) \\ \eta(y,z) \\ \eta(x,z) \end{bmatrix} > 0$$

$$Bx = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \eta(x,y) \\ \eta(y,z) \\ \eta(x,z) \end{bmatrix} \geq 0$$

$$Cx = \begin{bmatrix} a & c & -b \end{bmatrix} \begin{bmatrix} \eta(x,y) \\ \eta(y,z) \\ \eta(x,z) \end{bmatrix} = 0$$

From the dual system we obtain:

$$\gamma_1 = \frac{\mu_1 + \pi_1}{a}; \quad \gamma_1 = \frac{\mu_1 + \pi_2}{c}; \quad \gamma_1 = \frac{\mu_1 - \pi_3}{b}$$

Hence we obtain three equations:

$$[c-a]\mu_1 = -c\pi_1 + a\pi_2 \qquad [b-a]\mu_1 = -b\pi_1 - a\pi_3 \qquad [b-c]\mu_1 = -b\pi_2 - c\pi_3$$

Note that the RHS of the last two equations is weakly negative since $\pi \geq 0$ and $a, b, c > 0$. By Moderate Stochastic Transitivity if either $b > a$ or $b > c$, the LHS of one of the last two equations is positive and the system has no solution. Suppose $b = c$ and $a > b$ then $\pi_2 = \pi_3 = 0$ and by definition $\pi_1 > 0$ but set $\mu_1 > 0$ appropriately and the system has solution. If $b = c$ and $b > a$ then $\pi_2 = \pi_3 = 0$, $\pi_1 > 0$ and since $b, c > 0$ the system has no solution. Finally, if $b = c$ and $b = a$ then $\pi_1 = \pi_2 = \pi_3 = 0$ violating semi-positivity and the system has no solutions.

Finally, note that if $b = a = c = k$ then by FOCs we have:

$$u(x) - u(y) = \eta(x,y)k$$

$$u(y) - u(z) = \eta(y,z)k$$

$$u(x) - u(z) = \eta(x,z)k$$

since $u(x) - u(z) = u(x) - u(y) + u(y) - u(z)$:

$$\frac{u(x) - u(y)}{\eta(x,z)} + \frac{u(y) - u(z)}{\eta(x,z)} = \frac{u(x) - u(y)}{\eta(x,y)}$$

substituting we obtain

$$\frac{\eta(x,y)k}{\eta(x,z)} + \frac{\eta(y,z)k}{\eta(x,z)} = \frac{\eta(x,y)k}{\eta(x,y)}$$

and so

$$\eta(x,y) + \eta(y,z) = \eta(x,z)$$

.

$\square$

Proof. - (ii)

The following is a counterexample with $|A| = 4$ and the cost function being the Shannon Entropy.

$p(x,y) = p(z,w) = 0.75$

$p(x,z) = p(y,w) = 0.99$

$p(x,w) = 0.76$

$p(y,z) = 0.6$

MST* is satisfied since:

$p(x,z) > \min[p(x,y), p(y,z)]$

$p(y,w) > \min[p(y,z), p(z,w)]$

$p(x,w) > \min[p(x,y), p(y,w)]$

$p(x,w) > \min[p(x,z), p(z,w)]$

Solving the model we obtain, among all, the following triangle inequalities:

$$[u(x) - u(z)]/4.59511 + [u(z) - u(w)]/1.0986 > [u(x) - u(w)]/1.15267$$

$$[u(x) - u(y)]/1.0986 + [u(y) - u(w)]/4.59511 > [u(x) - u(w)]/1.15267$$

Since $u(x)$ and $u(w)$ are the max and the min, let $u(x) = 100$ and $u(w) = 0$.

$$[100 - u(z)]/4.59511 + [u(z)]/1.0986 > [100]/1.15267$$

$$[100 - u(y)]/1.0986 + [u(y)]/4.59511 > [100]/1.15267$$

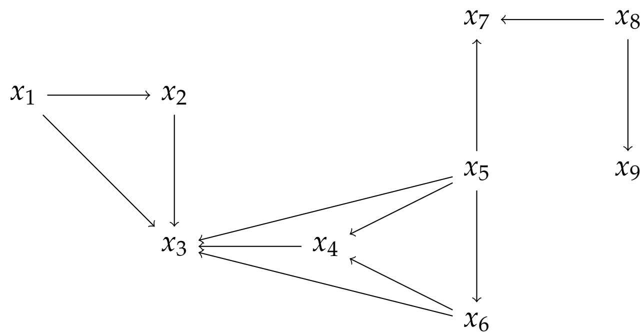
By the first inequality we have that $u(z)$ must be very close to 100 ($u(z) > 93.835...$). By the second inequality $u(y)$ must be very close to zero ($u(y) < 6.165...$). However since $p(y,z) = 0.6$ we must have that $u(y) > u(z)$ proving the contradiction.

$\square$

C.4 Examples

C.4.1 Deterministic Examples

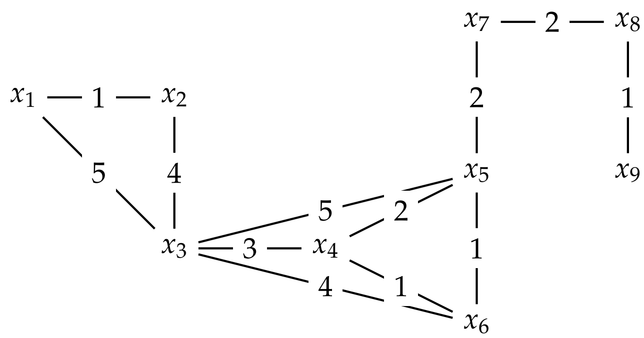
This is an example that applies the constructive proof of Theorem 1 to provide a BT representation of a transitive binary relation. Consider the following graph of a transitive relation:



The resulting matrix of utilities, where for all $x \in A$, $u(x) = |\{t \in A : t \not\succ x\}|$ is:

$$\left[\begin{array}{c|ccccccccc} & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \\ u & 8 & 7 & 3 & 6 & 8 & 7 & 6 & 8 & 7 \end{array}\right]$$

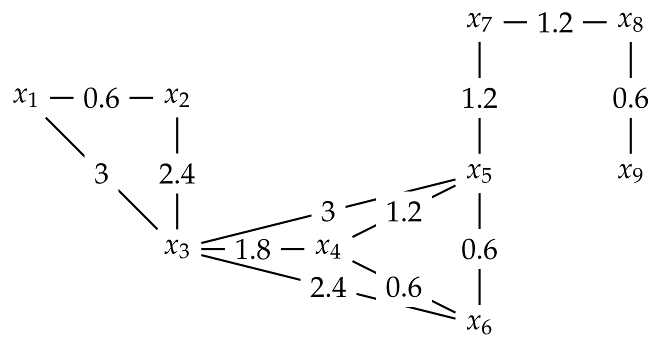
The resulting weighted graph (weights are difference in utilities):



Following Step 5 of the proof of Theorem 1, we solve for

$$\gamma^* = \underset{x,y \in A:[x \not\succ y \wedge y \not\succ x]}{\operatorname{argmax}} \frac{u(x) - u(y)}{\delta(x,y)}$$

where $\delta(x,y)$ is the minimum weighted path between x,y. The maximum is reached between x_8 and x_3 with $\gamma^* = \frac{5}{9}$. Hence, if we multiply the above weights for any $\gamma \in (\frac{5}{9},1)$ we can represent the partial order as a BT representation. For instance, let's $\gamma = 0.6$ the following graph arises. The reader may confirm that the binary relation is in fact represented.



C.4.2 Stochastic Examples

These are two examples that apply the constructive proof of Theorem 2 and 4. In particular, we start with two stochastic choice functions and construct a utility function, which is common to the two proofs, and then a function η that is a metric in the first case, while it violates Triangle Inequality in the second at exactly those triples that violate Strong Stochastic Transitivity.

Example 1. *We have the following stochastic choice function p:*

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ p(A) & (0.8,0.2) & (0.7,0.3) & (0.8,0.2) & (0.8,0.2) & (0.85,0.15) & (0.9,0.1) \end{array}\right]$$

- *The utilities are* $u(x) = 4$, $u(y) = 3$, $u(z) = 2$, $u(w) = 1$;
- *The set of binary sets is partitioned in four subsets:* $Y_1 = \{\{z,w\}\}$, $Y_2 = \{\{y,w\}\}$, $Y_3 = \{\{x,y\},\{y,z\},\{x,w\}\}$ *and* $Y_4 = \{\{x,z\}\}$;

Theorem 2 - Construction

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ \eta(A) & 3 & 18 & 9 & 3 & 2 & 0 \end{array}\right]$$

Note that Moderate Stochastic Transitivity is violated in the triple (x,y,z) and (x,y,w):*

$\eta(x,y)+\eta(y,z)-\eta(x,z)=-12$ $\qquad$ $\eta(x,y)+\eta(y,w)-\eta(x,w)=-4$

$\eta(x,z)+\eta(z,w)-\eta(x,w)=9$ $\qquad$ $\eta(y,z)+\eta(z,w)-\eta(y,w)=1$

Theorem 4 - Construction

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ \eta(A) & 2.\overline{33} & 5 & 7 & 2.\overline{33} & 4 & 1 \end{array}\right]$$

Note that Strong Stochastic Transitivity is always violated:*

$\eta(x,y)+\eta(y,z)-\eta(x,z)=-0.\overline{33}$ $\qquad$ $\eta(x,y)+\eta(y,w)-\eta(x,w)=-0.\overline{66}$

$\eta(x,z)+\eta(z,w)-\eta(x,w)=-1$ $\qquad$ $\eta(y,z)+\eta(z,w)-\eta(y,w)=-0.\overline{66}$

Example 2. *We have the following stochastic choice function p:*

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ p(A) & (0.75,0.25) & (0.99,0.01) & (0.76,0.24) & (0.6,0.4) & (0.99,0.01) & (0.75,0.25) \end{array}\right]$$

- *The utilities are $u(x)=4$, $u(y)=3$, $u(z)=2$, $u(w)=1$;*
- *The set of binary sets is partitioned in four subsets: $Y_1=\{\{x,z\},\{y,w\}\}$, $Y_2=\{\{x,w\}\}$, $Y_3=\{\{x,y\},\{z,w\}\}$ and $Y_4=\{\{y,z\}\}$;*

Theorem 2 - Construction

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ \eta(A) & 3 & 0 & 3 & 9 & 0 & 3 \end{array}\right]$$

Note that Moderate Stochastic Transitivity is always satisfied:*

$$\eta(x,y)+\eta(y,z)-\eta(x,z)=12 \qquad \eta(x,y)+\eta(y,w)-\eta(x,w)=0$$

$$\eta(x,z)+\eta(z,w)-\eta(x,w)=0 \qquad \eta(y,z)+\eta(z,w)-\eta(y,w)=12$$

Theorem 4 - Construction

$$\left[\begin{array}{c|cccccc} A & \{x,y\} & \{x,z\} & \{x,w\} & \{y,z\} & \{y,w\} & \{z,w\} \\ \eta(A) & 2.\overline{33} & 2 & 6 & 2.5 & 2 & 2.\overline{33} \end{array}\right]$$

Note that Strong Stochastic Transitivity is violated in the triple (x,y,w) and (x,z,w):*

$$\eta(x,y)+\eta(y,z)-\eta(x,z)=2.8\overline{33} \qquad \eta(x,y)+\eta(y,w)-\eta(x,w)=-1,\overline{66}$$

$$\eta(x,z)+\eta(z,w)-\eta(x,w)=-1,\overline{66} \qquad \eta(y,z)+\eta(z,w)-\eta(y,w)=2.8\overline{33}$$

C.5 Tversky & Russo (1969)

In Section 3.4.1 in Chapter 3, we connected Additive Perturbed Utility models to Fechnerian models. Here, we treat the case of Tversky & Russo (1969) model:

$$p(x,y)=F[u(x),u(y)]$$

where F is strictly increasing in $u(x)$ and strictly decreasing in $u(y)$. This model is completely characterized by a property called (Strict) Strong Stochastic Transitivity:

Definition. *A stochastic choice rule p satisfies (Strict) Strong Stochastic Transitivity [SSST] if for all $x,y,z \in A$:*

$$p(x,y)>\frac{1}{2} \;\&\; p(y,z)>\frac{1}{2} \quad\Rightarrow\quad p(x,z)>\max[p(x,y),p(y,z)]$$

Note that this property is the strongest introduced so far but it is similar to Strong Stochastic Transitivity*. Even if it is not directly connected with Additive Perturbed Utility models, a simple condition on η guarantees this model to be satisfied. The proof is straightforward. It is nonetheless interesting the counterexample that shows how the restriction imposed on η is stronger than the model of Tversky & Russo (1969).

Proposition 1. *Let a stochastic choice rule p has a BAPU representation. Take $x,y,z \in A$ such that $u(x) > u(y) > u(z)$; there exists a function η such that $\eta(x,y), \eta(y,z) \geq \eta(x,z)$ only if at $x,y,z \in A$, p satisfies (Strict) Strong Stochastic Transitivity.*

Proof. By $u(x) > u(y) > u(z)$ it must be:

$$\eta(x,z) \cdot [c'(p(x,z)) - c'(p(z,x))] > \eta(x,y) \cdot [c'(p(x,y)) - c'(p(y,x))]$$

and by $\eta(x,z) \leq \eta(x,y)$:

$$c'(p(x,z)) - c'(p(z,x)) > c'(p(x,y)) - c'(p(y,x))$$

that gives $p(x,z) > p(x,y)$.

Note that if, by contradiction, $p(x,z) = p(x,y)$ then by FOCs we have $u(x) - u(y) = \eta(x,y)[c'(p(x,y))) - c'(p(y,x))] < \eta(x,z)[c'(p(x,z)) - c'(p(z,x))] = u(x) - u(z)$. Hence, $\eta(x,z) > \eta(y,z)$.

Conversely, let $c(p)$ be the Shannon Entropy and let $p(x,y) = p(y,z) = 0.69 < p(x,z) = 0.7$ so that SSST is satisfied. Then, $c'(p(x,y)) - c'(p(y,x)) = c'(p(y,z)) - c'(p(z,y)) = 0.8$ and $c'(p(x,z)) - c'(p(z,x)) = 0.84$ but then the following system should have a solution with $\eta(x,y), \eta(y,z) \geq \eta(x,z)$:

$$u(x) - u(y) = 0.8\eta(x,y)$$

$$u(x) - u(z) = 0.84\eta(x,z)$$

$$u(y) - u(z) = 0.8\eta(y,z)$$

This can be rewritten as

$$0.8[\eta(x,y) + \eta(y,z)] = 0.84\eta(x,z)$$

let $\eta(x,y) = \eta(x,z) + \alpha_1$ and $\eta(y,z) = \eta(x,z) + \alpha_2$ for some $\alpha_1, \alpha_2 \geq 0$ then:

$$\frac{0.8}{0.84} = \frac{\eta(x,z)}{2\eta(x,z) + \alpha_1 + \alpha_2}$$

$$(1.6 - 0.84)\eta(x,z) = -0.8(\alpha_1 + \alpha_2)$$

that has clearly no solution for $\alpha_1, \alpha_2 \geq 0$ and $\eta(x,z) > 0$.

$\square$

References

Afriat, S. (1972). Efficiency estimation of production function. *International Economic Review*, *13*(3), 568–598.

Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, *125*(1), 40–68.

Aleskerov, F., Bouyssou, D., & Monjardet, B. (2007). Utility maximization, choice and preference.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutstrom, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583–618.

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengstrom, E. (2016). Risk aversion relates to cognitive ability: preferences or noise? *Journal of European Economic Association*, *14*(5), 1129–1154.

Andreoni, J., Gillen, B. J., & Harbaugh, W. T. (2013). The power of revealed preference tests: ex-post evaluation of experimental design. *Working paper*.

Apesteguia, J., & Ballester, M. A. (2015). A measure of rationality and welfare. *Journal of Political Economy*, *6*(123), 1278–1310.

Apesteguia, J., Ballester, M. A., & Lu, J. (2017). Single-crossing random utility models. *Econometrica*, *85*(2), 661–674.

Arrow, K. J. (1959). Rational choice functions and orderings. *Economica*, *26*(102), 121–127.

Badescu, B., & Weiss, W. (1987). Reflection of transitive and intransitive preferences:

a test of prospect theory. *Organizational Behavior and Human Decision Processes*, (39), 184–202.

Barberá, S., & Neme, A. (2017). Ordinal relative satisficing behavior. *Working paper*.

Beatty, T. K., & Crawford, I. A. (2011). How demanding is the revealed preference approach to demand? *American Economic Review*, *101*(6), 2782–2795.

Becker, G. S. (1962). Irrational behaviour and economic theory. *Journal of Political Economy*, (70), 1–13.

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

Bernheim, B. D., & Rangel, A. (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *The Quartely Journal of Economics*, *124*(1), 51–104.

Bombardini, M., & Trebbi, F. (2012). Risk aversion and expected utility theory: an experiment with large and small stakes. *Journal of the European Economic Association*, *10*(6), 1348–1399.

Bouacida, E., & Martin, D. (2020). Predictive power in behavioral welfare economics. *Journal of European Economic Association*.

Brady, R. L., & Rehbeck, J. (2016). Menu-dependent stochastic feasibility. *Econometrica*, *84*(3), 1203–1223.

Bronars, S. G. (1987). The power of nonparametric tests of preference maximization. *Econometrica*, *55*(3), 693–698.

Caliari, D. (2020). Consistency of choice in time and risk preferences. *Working paper*.

Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention and costly information acquisition. *American Economic Review*, *105*(7), 2183–2203.

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

Cattaneo, M. D., Ma, X., Masatlioglu, Y., & Suleymanov, E. (2018). A random attention model. *Working paper*.

Cavagnaro, D. R., & Davis-Stober, C. P. (2014). Transitive in our preferences, but transitive in different ways: an analysis of choice variability. *Decision*, *1*(2), 102–122.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., & Riella, G. (2019). Deliberately stochastic. *American Economic Review*, *7*(109), 2425–2445.

Chabris, C. C., Laibson, D., Morris, C. L., Schuldt, J. P., & Taubinsky, D. (2009). The allocation of time in decision making. *Journal of the European Economic Association*, (7), 628–637.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: between-subject and within-subject design. *Journal of Economic Behaviour & Organization*, (81), 1–8.

Danan, E., & Ziegelmeyer, A. (2006). Are preferences complete? an experimental measurement of indecisiveness under risk.

Dantzig, G. (1963). Linear programming and extensions. (Princeton).

Dasgupta, I., & Pattanaik, P. K. (2007). "regular" choice and the weak axiom of stochastic revealed preference. *Economic Theory*, (31), 35–50.

Day, B., Bateman, I., Carson, R., Dupont, D., Louviere, J., Morimoto, S., Scarpa, R., & Wang, P. (1987). Ordering effects and choice set awareness in repeat-response stated preference studies. *Journal of Environmental Economics and Management*, *1*(63), 73–91.

Dean, M., & Martin, D. (2015). Measure of rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics*, *3*(98), 524–534.

Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica*, *26*(3), 440–444.

Debreu, G. (1960). Review of r.d. luce, individual choice behaviour: a theoretical analysis. *American Economic Review*, (50), 186–188.

Dekel, E., Lipman, B. L., & Rustichini, A. (2001). Representing preferences with a unique subjective state space. *Econometrica*, *69*(4), 891–934.

Dutta, R. (2020). Gradual pairwise comparison and stochastic choice. *Working paper*.

Dziewulski, P. (2018). Just-noticeable difference as a behavioural foundation of the critical cost-efficiency index. *Working paper*.

Echenique, F., Lee, S., & Shum, M. (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy*, *119*(6), 1201–1223.

Echenique, F., Saito, K., & Tserenjigmid, G. (2018). The perception-adjusted luce model. *Mathematical Social Sciences*, *93*, 67–76.

Enke, B., Gneezy, U., Hall, B., Martin, D. C., Nelidov, V., Offerman, T., & van de Ven, J. (2021). Cognitive biases: mistakes or missing stakes? *NBER Working paper 28650*.

Famulari, M. (1995). A household-based, nonparamtetric test of demand theory. *The Review of Economics and Statistics*, (77), 372–83.

Fishburn, P. C. (1970). Intransitive indifference with unequal indifference intervals. *Journal of Mathematical Psychology*, (7), 144–149.

Fishburn, P. C. (1973). Binary choice probabilities: on the varieties of stochastic transitivity. *Journal of Mathematical Psychology*, (10), 327–352.

Fishburn, P. C. (1978). Choice probabilities and choice functions. *Journal of Mathematical Psychology*, (18), 205–219.

Fishburn, P. C. (1999). Preference structures and their numerical representations. *Theoretical Computer Science*, (217), 359–383.

Fishburn, P. C., & Rubinstein, A. (1982). Time preference. *International economic review*, *23*(3), 677–694.

Fosgerau, M., Melo, E., Palma, A. D., & Shum, M. (2017). Discrete choice and rational inattention: a general equivalence result. *Working paper*.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Frick, M. (2016). Monotone threshold representations. *Theoretical Economics*, (11), 757–772.

Fudenberg, D., Iijima, R., & Strzalecky, T. (2014). Stochastic choice and revealed perturbed utility. *Report*.

Fudenberg, D., Iijima, R., & Strzalecky, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica*, *83*(6), 2371–2409.

Fudenberg, D., Kleinberg, J., Liang, A., & Mullainathan, S. (2019). Measuring the completeness of theories. *PIER Working Paper No. 18-010*.

Goodin, R. E., & List, C. (2006). Special majorities rationalized. *British Journal of Political Science*, (36), 213–241.

Gul, F., Natenzon, P., & Pesendorfer, W. (2014). Random choice as behavioral optimization. *Econometrica*, *82*(5), 1873–1912.

Gul, F., & Pesendorfer, W. (2006). Random expected utility. *Econometrica*, *74*(1), 121–146.

Harbarugh, W. T., Krause, K., & Berry, T. R. (2001). Garp for kids: on the development of rational choice behavior. *American Economic Review*, *91*(5), 1539–1545.

Harrison, G. W., & Rutstrom, E. E. (2008). Risk aversion in the laboratory. *in James C. Cox, Glenn W. Harrison (ed.) Risk Aversion in Experiments*, (pp. 41–196).

Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: the effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology & Marketing*, *26*(3), 204–212.

He, J., & Natenzon, P. (2018). Moderate expected utility. *Working paper*.

Hey, J. (2001). Does repetition improve consistency? *Experimental economics*, *4*(1), 5–54.

Hey, J., & Carbone, E. (1995). Stochastic choice with deterministic preferences: an experimental investigation. *Economics Letters*, *47*(2), 161–167.

Horan, S., & Sprumont, Y. (2016). Welfare criteria from choice: An axiomatic analysis. *Games and Economic Behavior*, (99), 56–70.

Houtman, M., & Maks, J. A. H. (1985). Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve Methoden*, (19), 89–104.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *The Journal of Consumer Research*, *9*(1), 90–98.

Iyengar, S. S., & Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, (94), 530–539.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: can one desire too much of a good thing? *Journal of personality and social psychology*, *79*(6), 995–1006.

Johnsonbaugh, R. (1998). A discrete intermediate value theorem. *College Mathematics Journal*.

Kalai, G., Rubinstein, A., & Spiegler, R. (2002). Rationalizing choice functions by multiple rationales. *Econometrica*, *70*(6), 2481–2488.

Kim, Y.-I., & Lee, J. (2012). Estimating risk aversion using individual-level survey data. *The Korean Economic Review*, *28*(2), 221–239.

Klamler, C. (2008). A distance measure for choice functions. *Social Choice Welfare*, (30), 419–425.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). Foundations of measurement. *Volume I*.

Ladenburg, J., & Olsen, S. B. (2008). Gender-specific starting point bias in choice experiments: evidence from an empirical study. *Journal of Environmental Economics and Management*, *3*(56), 275–285.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quartely Journal of Economics*, *112*(2), 443–477.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2017). When more is less: limited consideration. *Journal of Economic Theory*, (170), 70–85.

Luce, D. R. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, *24*(2), 178–191.

Luce, D. R. (1959). Individual choice behavior: a theoretical analysis.

Machina, M. (1985). Stochastic choice functions generated from deterministic preferences over lotteries. *The Economic Journal*, (95), 575–594.

Mandler, M., Manzini, P., & Mariotti, M. (2012). One million answers to twenty questions: choosing by checklists. *Journal of Economic Theory*, (147), 71–92.

Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review*, *97*(5), 1824–1839.

Manzini, P., & Mariotti, M. (2014). Stochastic choce and consideration sets. *Econometrica*, *82*, 1153–1176.

Manzini, P., & Mariotti, M. (2018). Dual random utility maximization. *Journal of Economic Theory*, (177), 162–182.

Manzini, P., Mariotti, M., & Mittone, L. (2010). Choice over sequences of outcomes: theory and experimental evidence. *Theory and Decision*, (69), 327–354.

Marschak, J., & Block, H. (1960). Random orderings and stochastic theories of responses. *Contributions to Probability and Statistics*, (Stanford University Press).

Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, *102*(5), 2183–2205.

Matejka, F., & McKay, A. (2015). Rational inattention to discrete choices: a new foundation for the multinomial logit. *American Economic Review*, *105*(1), 272–298.

Mattsson, L. G., & Weibull, J. (2002). Probabilistic choice and procedurally bounded rationality. *Games and Economic Behaviour*, (41), 61–78.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. *Frontiers of Econometrics*, (pp. 105–142).

McFadden, D. (1980). Econometric models for probabilistic choice among products. *The Journal of Business*, *53*(3), S13–S29.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, *10*(1), 6–38.

Meyer, B. D., & Mongin, P. (1995). A note on affine aggregation. *Economics Letters*, *47*(2), 177–183.

Monjardet, B. (1980). Metrics on partially ordered posets - a survey. *Discrete Mathematics*, (35), 173–184.

Nakamura, Y. (2002). Semimetric thresholds for finite posets. *Mathematical Social Sciences*, (44), 37–43.

Natenzon, P. (2019). Random choice and learning. *Journal of Political Economy*, *127*(1), 419–457.

Nishimura, H. (2017). The transitive core: inference of welfare from nontransitive preference relation. *mimeo*.

Ok, E., & Nishimura, H. (2018). Preference structures. *Working paper*.

Ok, E., & Tserenjigmid, G. (2019). Deterministic rationality of stochastic choice behavior. *Working paper*.

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

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

Plott, C. R. (1993). Rational individual behavior in markets and social choice processes. *Working paper - California Institute of Technology*.

Ravid, D., & Stevenson, K. (2019). Focus, then compare. *Working paper*.

Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: an eye tracking study. *American Economic Review*, (101), 900–926.

Rubinstein, A. (1980). Ranking the participants in a tournament. *SIAM Journal on Applied Mathematics*, *38*(1), 108–111.

Rubinstein, A., & Salant, Y. (2012). Eliciting welfare preferences from behavioural data sets. *The Review of Economic Studies*, (79), 375–387.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences*, *21*(2), 153–167.

Sen, A. (1969). Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies*, *36*(3), 381–393.

Sen, A. (1971). Choice functions and revealed preference. *The Review of Economic Studies*, *38*(3), 307–317.

Sippel, R. (1997). An experiment on the pure theory of consumer's behaviour. *Economic Journal*, *107*(444), 1431–1444.

Soltani, A., Martino, B. D., & Camerer, C. (2012). A range-normalization model of context-dependent choice: a new model and evidence. *PLoS computational biology*, *7*(8).

Sopher, B., & Narramore, M. J. (2000). Stochastic choice and consistency in decision making under risk: an experimental study. *Theory and Decision*, *48*(4), 323–350.

Swofford, J. L., & Whitney, G. A. (1987). Nonparametric test of utility maximization and weak separability for consumption, leisure and money. *The Review of Economics and Statistics*, (69), 458–64.

Thurstone, L. (1927). A law of comparative judgement. *Psychological Review*, *34*(4), 272–286.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, (76), 31–48.

Tversky, A. (1972). Elimination by aspects: a theory of choice. *Psychological Review*, *4*(79), 281–299.

Tversky, A., & Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, *6*(1), 1–12.

van den Brink, R., & Gilles, R. P. (2003). Ranking by outdegree for directed graphs. *Discrete Mathematics*, (271), 261–270.