



HAL
open science

Choices, Preferences, and Welfare

Elias Bouacida

► **To cite this version:**

Elias Bouacida. Choices, Preferences, and Welfare. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2019. English. NNT: 2019PA01E017 . tel-02521809

HAL Id: tel-02521809

<https://theses.hal.science/tel-02521809>

Submitted on 27 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Paris I Panthéon Sorbonne
UFR d'Économie

Laboratoire de rattachement: Paris Jourdan Sciences Economiques

THÈSE

Pour l'obtention du titre de Docteur en Sciences Economiques

Présentée et soutenue publiquement

le 10 juillet 2019 par

Elias Bouacida

Choices, Preferences, and Welfare

Sous les directions de :

Jean-Marc Tallon, Directeur de recherche CNRS, PjSE, Professeur à l'école d'économie de Paris
Daniel Martin, Associate Professor, Kellogg School of Management, Northwestern University

Jury :

Rapporteurs :

Eric Danan, Chargé de recherche CNRS – HDR, THEMA, Université de Cergy-Pontoise
Georgios Gerasimou, Senior Lecturer, Université de Saint-Andrews

Examineurs :

Olivier l'Haridon, Professeur à l'université de Rennes
Stéphane Zuber, Directeur de recherche CNRS, CES, professeur à l'école d'économie de Paris

Contents

Remerciements	9
1 Introduction	11
1.1 Revealed Preferences	14
1.2 Beyond Classical Preferences	20
1.2.1 Relaxing Transitivity and Menu-Independence	22
1.2.1.1 Partial Order	24
1.2.1.2 Interval Order	25
1.2.1.3 Semi-Order	25
1.2.1.4 Monotone Threshold	26
1.2.1.5 Menu-Dependent Threshold	26
1.2.1.6 Context-Dependent Threshold	27
1.2.2 Robust Revealed Preferences	27
1.2.2.1 The Strict Unambiguous Choice Relation	29
1.2.2.2 The Transitive Core	30
1.2.2.3 Other Welfare Relations	30
1.3 Organization of the Dissertation	31
2 Predictive Power in Behavioral Welfare Economics	33
2.1 Introduction	33
2.1.1 Empirical Findings in Behavioral Welfare Economics	37
2.2 Data	38
2.2.1 Experimental Data	39

2.2.2	Consumption Data	40
2.2.2.1	Analysis Sample: Panelists	40
2.2.2.2	Analysis Sample: Demographic Characteristics	41
2.2.2.3	Analysis Sample: Bundles	42
2.2.2.4	Analysis Sample: Prices	43
2.2.2.5	Additional Considerations	45
2.3	Results	45
2.3.1	Inconsistencies in Revealed Preferences	45
2.3.2	Inconsistencies in SUCR and TC	46
2.3.3	Completeness of SUCR and TC	47
2.3.4	Predictive Power of SUCR and TC	49
2.3.5	Predictive Power and Revealed Preference Properties	52
2.3.5.1	Number of Direct RP Cycles	54
2.3.5.2	Fraction of RP Cycles that are Direct	55
2.3.5.3	Regression Analysis	56
2.4	Discussion and Conclusion	56
3	Identifying Choice Correspondences	59
3.1	Introduction	59
3.2	Definition of the Pay-For-Certainty Method	63
3.2.1	General Setup	63
3.2.2	Definition of the <i>Pay-For-Certainty</i> Method	64
3.2.2.1	Incentive to Choose Several Alternatives	65
3.2.2.2	Selection Mechanisms	66
3.2.2.3	Linear versus Proportional Bonus Payment	68
3.2.3	Pay-For-Certainty with a Uniform Selection Mechanism	68
3.3	Identification of the Choice Correspondence	69
3.3.1	Assumptions	69
3.3.2	Identification of the Choice Correspondence	72

<i>CONTENTS</i>	5
3.3.2.1	Partial Identification of the Choice Correspondence 72
3.3.2.2	Full Identification of the Choice Correspondence 74
3.3.3	Consistency when Full Identification Fails 75
3.3.3.1	Compatibility with Classical Preferences 76
3.3.3.2	Compatibility with a Partial Order 79
3.3.3.3	Compatibility with Menu-Dependent Threshold 81
3.3.3.4	Compatibility with Fixed Point 82
3.4	Conclusion 82
4	<i>Pay-for-certainty in an Experiment</i> 85
4.1	Introduction 85
4.1.1	Menu Choice in the Literature 88
4.2	Design of the Experiment 89
4.2.1	Tasks 90
4.2.2	Timing 91
4.2.3	Information Treatments 92
4.2.4	Choices 92
4.2.5	Questionnaire 94
4.2.6	Data 94
4.2.7	Demographics Characteristics 94
4.3	Comparing Three Revelations Method 95
4.3.1	The Benchmark: Forced Single Choice 95
4.3.2	0-Correspondences 97
4.3.3	1-Correspondences 98
4.3.4	Identified Choice Correspondence 99
4.3.4.1	Fully Identified Choice Correspondences 100
4.3.4.2	Partially Identified Choice Correspondences 101
4.3.5	Related Literature and Discussion 101
4.4	Going Further with Correspondences 102

4.4.1	Intransitive Indifference	103
4.4.2	Menu-Dependent Choices	104
4.4.3	Discussion	104
4.5	Influence of the Information Provided	105
4.6	Conclusion	109
5	Conclusion	111
A	Appendix of Chapter 2	113
A.1	Choice Set Size	113
A.2	Correlations for Uniform Random Demands	113
A.2.1	Uniform Random Methodology: Consumption Data	113
A.2.2	Uniform Random Methodology: Experimental Data	114
A.2.3	Results	114
B	Appendix of Chapter 3	117
B.1	Own Randomization	117
B.2	Extending Pay-For-Certainty to Infinite Sets	119
B.2.1	Infinite Choice Sets	119
B.2.2	Finite Choice Sets on Infinite Sets	120
C	Appendix of Chapter 4	121
C.1	Design	121
C.1.1	Tasks	121
C.1.2	Instructions	124
C.1.2.1	Gains in the experiment	124
C.1.2.2	Remarks	124
C.1.2.3	Instructions	125
C.2	Explanatory Power of Different Axioms	126
C.3	Size of the Chosen Sets	127
C.4	Weak Revealed Preferences	129

C.4.1	With Forced Single Choice	129
C.4.2	0-Correspondences	130
C.5	Results with High Payments	131
C.6	Choice Functions and Choice Correspondences on the Same Subjects	133
C.7	Aggregating Choices	134
C.7.1	Information and Maximal Alternatives	134
C.7.2	An Alternative Determination of Maximal Alternatives	135
D	Analytical Computer Programs	139
D.1	Chapter 2	139
D.2	Chapter 4	139
	Bibliography	139
E	Résumé en Français	153
E.1	Introduction	153
E.2	Pouvoir prédictif en économie du bien-être comportementale	157
E.3	Identification des correspondances de choix	162
E.4	Payer pour la certitude dans une expérience	166
E.5	Conclusion	169

Remerciements

I am thankful to my colleagues, my family, and my friends, who supported me and helped me for the last five years and beyond! Without you, this Ph.D. would not look the same.

Je tiens à remercier mes collègues, ma famille et mes amis, qui m'ont aidé et soutenu tout au long de ces cinq dernières années, et bien plus ! Sans vous, cette thèse ne sera pas ce qu'elle est.

I want to thank Daniel Martin, who lead my first steps in research. You were and still are crucial support, and I have really enjoyed working under your direction, and now with you. You are also the one who recommended me to Jean-Marc Tallon, which may be the single most significant result of my master thesis dissertation.

Jean-Marc Tallon, merci ! Merci pour ta porte toujours ouverte, ta disponibilité et ton soutien sans faille, ta gentillesse, et surtout, tes commentaires toujours pertinents, même si j'ai parfois mis du temps à les comprendre... Cette thèse te doit aussi beaucoup et ces cinq années auraient été très différentes sans toi.

J'aimerais aussi remercier Stéphane Zuber, qui a suivi cette thèse de près. Ta rigueur et tes commentaires ont toujours été pertinents et m'ont toujours éclairé.

I would like to thank the members of my jury, who accepted to spend time on my thesis and have already given me great comments. Eric Danan, in particular, was available way before the jury took shape, and talking with you was always very fruitful, at the MSE, Cergy or PSE.

To all the researchers I had the chance to talk with during my Ph.D., at the MSE, PSE, Northwestern, and Nanterre in particular, thank you! Nicolas Jacquemet and Béatrice Boulu-Reshef's advice, as well as all the workgroup in experiments, have been of great help when I started building my experiment, and well after. Maxim Frolov is an excellent and dedicated lab' programmer, without him, I would have lost a ton of time running my first experiment. I have participated in many conferences and seminars during my Ph.D., thanks to the administrative support at MSE and PSE for making the process smoother. All the participants in these conferences make the process of research more collaborative and more enjoyable.

On that note, I would like to thank all the researchers, Ph.D. and administrative staff involved in the job market training, without whom I would not have gotten the job at Lancaster. Véronique

Guillot in particular, who was and is always available for Ph.D. students, and who shared the good and the bad news of all of us. Eyal Winter and Maurizio Zanardi allowing me to pursue my research interest in the next three years at Lancaster University. I am eager to come!

The discussion, support, and presence of my colleagues and friends were all integral and essential part of the last five years. To name a few and forgetting many more. My officemates: Thomas, Juliette, Aya, Joanna, Ariane, Emanuela, Sasha, Sulin, Cem, Annali, Alexia, Maiting, Avner, Ezgi, Ole, Florian, Peio, Zouhair, Anthony, Dalia, George, Sandrine, Julianna. My colleagues and lunch mates: Giulio, Guillaume, Shaden, Julien, Rémi, Léontine, Anna, Justine, Antoine, Antoine, Matthieu, Alexandra, Simon, Philippe, Juni. The two reading groups: Simon, Rémi, Avner, Philippe, Quentin, Anthony, and Niels. Paul and Juliette, the team of committed representatives. To the Synapse members over the years, and in particular, Julien, Zahra and Juni. A special thank for Zahra, who made my arrival and my life in Chicago much, much more relaxed and funnier! All of you, and many others, have made my Ph.D. more fruitful and more enjoyable, in one way or another, thank you!

Ma famille, mes proches et mes amis. Ces cinq dernières années n'ont pas toujours été faciles. Merci d'avoir été là, votre soutien a été, est et sera toujours essentiel. Merci à Quentin, Benjamin, Claire, Pierre, Maxime, Clément, Marion, Andéol, Johan, Roman, pour nos super week-ends, et ce n'est pas fini : votre prochaine destination, c'est Lancaster ! Merci à Lucile, nous avons commencé nos doctorats en même temps, et nous allons les finir presque en même temps ! Merci à Inès, Inès, Haïfa, Sabrina et Youssr, mes tunisiens en France, il y a des choses que vous seules comprenez. Grand merci à Alter'Actions, et en particulier, Isabelle, Michel, Benjamin, Francesca, Amandine, Nicolas et Audrey, vous m'avez rattaché à une réalité différente de la recherche cinq années durant, qui m'ont aidé à garder les pieds sur terre. Merci à une des parties les plus stable de ma vie jusqu'à présent, le karaté, et en particulier, Senseï, Léa, Laurant, Philippe, Carlos, Vuslat, Pierre, Lahcène, Jean, Catherine, et tous les autres. Vous m'avez aussi rattaché à un monde différent de la recherche et vous allez me manquer. Merci à mes colocs des 10 dernières années.

Ma famille, et en particulier mes cousins et cousines, vivant à Paris n'est pas et n'a pas été très nombreuse au cours de ces années, mais est toujours d'un grand soutien. Grâce à vous, je me sens plus chez moi à Paris ! Un remerciement particulier à Inès, qui a déjà vécu une thèse avant la sienne. Maintenant, je t'en dois une !

Enfin, un remerciement tout particulier à mes très proches, père, mère, sœurs et frère.

Chapter 1

Introduction

This dissertation studies the link between choices, preferences, and individual welfare. One object of interest in this dissertation is an *alternative*. An alternative can take various forms; it may be a bundle purchased from the supermarket, a meal, the possibility to register in a university, the date to which you will have dinner with your parents. *Choices* are made between different alternatives. *Preferences* are a ranking of the alternatives to choose from. *Welfare* is the well-being, or satisfaction, obtained from an alternative.

The assessment of welfare is an integral part of economics, in particular for the evaluation of public policy and policy recommendations. One method to study individual welfare is to identify the satisfaction decision makers retain from their choices. It can be a relative or an absolute satisfaction, and it can be obtained by directly asking, for instance, in subjective well-being studies, or by deducing it from the preferences of individual decision makers. Preferences are not directly observable, but they can be inferred. The principle we will use in this dissertation was introduced by Samuelson (1938)'s seminal paper; it is called *revealed preferences*. Revealed preferences link choices and preferences. Little (1949) was the first to propose to use revealed preferences for individual welfare analysis.

There is an apparent contradiction between Samuelson (1938)'s project with revealed preferences and Little (1949)'s use of revealed preferences for welfare analysis. Indeed, Samuelson (1938)'s analysis stated aim was (emphasis are ours):

I propose, therefore, that we start anew in direct attack upon the problem, dropping off the *last vestiges* of the *utility* analysis.

If indeed, revealed preferences aimed at dropping any trace of utility in economic analysis, it would not be suitable for welfare analysis, as the *revealed* preference would not represent anything meaningful from a welfare perspective. It would only be an *as if* tool to model choices made by decision

makers, a convenient mathematical representation to represent choices made. Little (1949)'s words (the theory of consumer behavior should be understood as the revealed preference theory):

In the theory of Consumer's Behaviour, as at present formulated, the preferences (in the subjective sense) which are discovered by asking questions are linked to market behaviour by the postulate that people do in fact try to maximize satisfaction. It is thus false to suppose that the theory, as it stands at present, is concerned only with choice. It is concerned also with likes and dislikes.

Little (1949) links choices with satisfaction. It can only be understood as decision makers have a utility or a preference, which they maximize. Revealed preferences try to recover this utility from the observed choices. In Amartya Sen (1973)'s words:

The rationale of the revealed preference approach lies in this assumption of revelation and not in doing away with the notion of underlying preferences, despite occasional noises to the contrary.

To summarize, using revealed preferences for welfare analysis requires two things. First, that decision makers have a preference (utility), which is the right guide to their welfare, and second, that revealed preferences identify it. To understand whether these assumptions are warranted, it is helpful to understand the core idea of revealed preferences. It is quite intuitive. If John chooses an apple instead of a banana when both are available, then he probably prefers apples to bananas. Thus, he is better off with an apple rather than a banana. In this simple example, we can go directly from choices to welfare. Preferences are needed because not all choices are observed. If we observe that John chooses apples over bananas, and bananas over oranges, and think in terms of welfare, we probably expect him to choose apples over oranges, even *without* observing a choice between apples and oranges.

While intuitively straightforward, the reasoning faces hurdles in practice. For instance, what should we conclude if John sometimes chooses bananas and sometimes apples, or if he chooses apples over bananas, bananas over oranges, and oranges over apples? It is harder to reveal a preference with a welfare interpretation in these cases. The second problem is that the preference built from observed choices might not be the one we should use for welfare analysis. What should we think of the choices of a smoker choosing to smoke, while knowing that it is unhealthy, for instance? The use of choices as a guide to welfare is has been debated for a long time, in different forms. Amartya Sen (1973), for instance, provides an early critique. One of the most recent iterations of this debate is in behavioral welfare economics, between, for instance, Bernheim and Rangel (2009) and Salant and Rubinstein (2008). We will come back to this last debate in Chapter 2. Mongin and d'Aspremont (1998) provide a normative defense of revealed preferences for welfare analysis.

If we accept that we can use choices as a guide to welfare, in the strong sense implied by revealed preferences, we still have a positive problem, like the lack of transitivity in practice. Koo (1963) provided the first example, Choi, Fisman, et al. (2007), Choi, Kariv, et al. (2014), Dean and Martin (2016), among others, provide more recent illustrations. In those cases, it is much less straightforward how to build a preference from observed choices, let alone think about welfare. Revealed preferences require some consistency between observed choices to be used in welfare analysis. In other words, choices must obey some rules so that we can build a preference. Indeed, if there are no rules to the observed choices, then there is no hope to build a model to explain the choices. Because the conditions implied by the rules are violated in practice, it means that the process we assume is not the right one.

There are two possible reactions to these failures. The first is to consider that the revealed preference approach is doomed and to try to bring insights on the decision-making processes to build a more accurate model of decision making. The second is to look at the hypotheses behind revealed preferences and the setup employed and to look at why revealed preferences seem to fail. The main objective of this dissertation is to improve the *revelation* of preferences, and we will take these two directions to do so.

First, the revealed preference approach assumes that decision makers can be modeled *as if* they maximize a preference. It is not how decisions are made in general, as research in psychology and behavioral economics have shown. So revealed preferences is not a good model of the human decision making processes. Instead, decision makers can be described as using heuristics to make choices, a fact we have known at least since Tversky and Kahneman (1974). This understanding led to the development of behavioral economics. Behavioral economics usually keeps the choice as relevant to welfare analysis, even knowing that the revealed preference might not be the *preference*, and indeed, the preference might not exist.

The study of welfare in behavioral economics was often made using *ad hoc* models of the heuristics used to make choices, in order to tease out the preference from noisy choices. This kind of approaches has been criticized because of their *ad hoc* nature. Behavioral welfare economics emerged in response, following the seminal papers of Bernheim and Rangel (2009) and Salant and Rubinstein (2008). This literature aims is to provide a more rigorous basis for the study of welfare when choices are not the results of utility maximization. So far, the literature on the topic has been primarily theoretical (Ambuehl, Bernheim, and Lusardi (2014), Bernheim, Fradkin, and Popov (2015) are two exceptions). This dissertation aims at providing empirical insights into the study of welfare in the presence of biases. In Chapter 2, with my co-author Daniel Martin, we study two propositions from the behavioral welfare economics literature.

Chapter 3 provides a method to study revealed preferences using a broader method compared to the empirical studies done so far by extending the domain of choice. It reveals some limits from the current practices of revealed preferences, and notably the possible underestimation of indifference

with the current methods. It might explain why decision makers seem sometimes mistaken in traditional experiments. Chapter 4 illustrates the method of Chapter 3 with an experiment.

Section 1.1 introduces in more details the revealed preference approach, as well as the conditions for welfare assessment. Section 1.2 goes beyond the classical revealed preference approach while keeping its central insight.

1.1 Revealed Preferences

Economists traditionally model decisions made by individuals as the result of preference maximization, or, often equivalently, utility maximization. A *utility* is a function from the set of alternatives to the real numbers, which implies a ranking over the alternatives. The *maximization* is the process used to determine the alternative chosen: it should be the highest ranked according to the preference or the alternative which yield the highest utility. This strong link between the *observed* choice and the *unobserved* preference suggest that we can recover the preference or utility from the observed choices. It is the idea behind *revealed preferences*. Amartya Sen (1997) discusses the weaknesses of this approach, which is still a backbone of economic analysis. We use some departure of the process of utility maximization in this dissertation, but we still assume that choices reflect preferences.

Indeed, the process of utility maximization yields some conditions on the choices. These conditions differ depending on the choice environment. The conditions on infinite data (typically, consumption data, i.e., choices with quantities and prices), were first given by Ville and Newman (1951), Houthakker (1950) and Afriat (1967)'s *Weak Axiom of Revealed Preferences*, and Hal R. Varian (1982)'s *General Axiom of Revealed Preferences*. Crawford and De Rock (2014) and Adams and Crawford (2015) are recent reviews on revealed preferences methods with consumption data and beyond. The conditions on finite data (typically, experimental data, i.e., the choice of items from a set of available alternatives, without a price), were first given by Arrow (1959) and Richter (1966). These conditions characterize the set of observed choices that are *rationalizable* by a preference, i.e., that can be reversed in order to build preference.

These conditions impose some consistency on observed and unobserved choices and have out-of-sample consequences, which is what leads us to say that John should choose apples over oranges earlier even without observed the choice between apples and oranges. Observed choices restrict future choices, by, for instance, restraining what kind of new bundle can be chosen within a given price line. One example of such restrictions was given by Hal R. Varian (1982). Two main measures can be used to assess the usefulness of a model in out-of-sample terms. First, *predictive power*, which is roughly a measure of how many alternatives could be chosen and be compatible with the preference. The higher the predictive power, the more precise is the model. Second, *predictive success*, which tells us whether the observed choice is compatible with the preference built on the

rest of the observed choices. Beatty and Crawford (2011) provide a recent example of the use of such out-of-sample measures of predictive power and predictive success. In Chapter 2, we use predictive power to assess the validity of two behavioral welfare economics measures that we will introduce in Section 1.2.2.

This dissertation studies revealed preferences on infinite and finite data. Formally, choices on finite data can be modeled by a *choice correspondence* or a *choice function*. Call X is the grand set of alternatives, $\mathcal{P}(X)$ the set of all non-empty finite subsets of X , i.e., $\mathcal{P}(X) = 2^X \setminus \emptyset$. $\mathcal{P}(X)$ is the set of all possible *choice sets*. x, y, z are alternatives (elements) of X .

Definition 1.1 (Choice Correspondence). A **choice correspondence** on $\mathcal{P}(X)$ associates to non-empty subsets of X a choice $c(S)$, which is a non-empty subset of S (and thus an element of $\mathcal{P}(X)$).

$$\begin{aligned} c : \mathcal{P}(X) &\rightarrow \mathcal{P}(X) \\ S &\rightarrow c(S) \subseteq S \end{aligned}$$

Notice, that we do not allow the choice correspondence to be empty valued that is, $c(S) \neq \emptyset$. It constrains choices over singletons: $c(\{x\}) = \{x\}$ for all $x \in X$. We will, as a consequence, omit choices over singletons in the whole dissertation.¹ *Choice functions* are a special kind of choice correspondences, where the chosen set $c(S)$ is made of exactly one alternative.

Definition 1.2 (Choice Function). A **choice function** on $\mathcal{P}(X)$ associates to non-empty subsets of X a chosen alternative x , which is an element of S .

$$\begin{aligned} c : \mathcal{P}(X) &\rightarrow X \\ S &\rightarrow x \in S \end{aligned}$$

Importantly, finite data does not have an underlying structure: no prices, no quantities, no ranking *a priori* on the alternatives. The latter is not a requirement, as studies of preferences over lotteries have an underlying structure which can at least partially rank the alternatives.

Definition 1.3 (Situation). A *situation* i is one observed choice. With finite data, it is a choice set S_i and a choice from this choice set, $c(S_i)$.

The observable with finite data is a collection of \mathcal{N} situations $(S_i, c(S_i))_{i \in \mathcal{N}}$.² Situations allow us to observe several choices from the same choice set.

The infinite data we care about here is *consumption data*. Consumption data relies on two primitives: for each situation $i \in \mathcal{N}$, a vector of prices p_i and a vector of quantities x_i . p_i and x_i are vectors

¹We could expand the definition of choice functions and choice correspondences to the empty set, by assuming that the choice in the empty set is the empty set, i.e., $c(\emptyset) = \emptyset$. It does not provide any additional insight.

² \mathcal{N} may be finite or infinite, *a priori*. In practice, however, it is necessarily finite.

of \mathbb{R}^m , where m is the number of available alternatives. The natural order on \mathbb{R}^m therefore ranks them. Observations with consumption data are a family of prices and quantities in \mathcal{N} situations: $(p_i, x_i)_{i \in \mathcal{N}}$.

In all the dissertation, R, P and I are binary relations on X , i.e., subsets of X^2 . We will note xRy for $(x, y) \in X^2$, and similarly, for P and I . We can define *revealed preferences* in these three different cases.

Definition 1.4 (Revealed Preferences). Revealed preferences are a collection of six binary relations, (R^0, P^0, I^0, R, P, I) , defined as:

- R^0 is the *directly revealed preference*.
- P^0 is the *strictly directly revealed preference*.
- I^0 is the *directly revealed indifference*.
- R is the *revealed preference*: x is *revealed preferred* to y , noted xRy , if there exists x_1, x_2, \dots, x_n such that $xR^0x_1, x_1R^0x_2, \dots, x_nR^0y$ (potentially with some R^0 being P^0 or I^0).
- P is the *strict revealed preference*: x is *strictly revealed preferred* to y , noted xPy , if there exists x_1, x_2, \dots, x_n such that $xR^0x_1, x_1R^0x_2, \dots, x_nR^0y$ and at least one of them is strict P^0 .
- I is the *revealed indifference*: x is *revealed indifferent* to y , noted xIy , if there exists x_1, x_2, \dots, x_n such that $xI^0x_1, x_1I^0x_2, \dots, x_nI^0y$.

The three last relations are the transitive closures of the first three.

In general, P is the *asymmetric* part of the revealed preference relation R , and I is its *symmetric* part. That is, xPy if and only if xRy and not yRx and xIy if and only if xRy and yRx . R^0, P^0 , and I^0 are mostly transitory tools. We are mostly interested in R, P , and I , which are supersets of R^0, P^0 , and I^0 . It means that R provides a summary of all the information in (R^0, P^0, I^0, R, P, I) , and is often called the *revealed preference*. We will sometimes denote R as \succeq with \succ being P and \sim being I . We can now turn to the definition of revealed preferences with finite and infinite data. We introduce two definitions of revealed preferences on finite data, the *strict* and the *weak* revealed preferences.

Definition 1.5 (Strict Revealed Preferences). With strict revealed preferences, we assume that chosen alternatives are strictly better than unchosen alternatives:

- xR^0y if and only if there exists $i \in \mathcal{N}$, $x \in c(S_i), y \in S_i$.
- xP^0y if and only if there exists $i \in \mathcal{N}$, $x \in c(S_i), y \in S_i \setminus c(S_i)$.
- xI^0y if and only if there exists $i \in \mathcal{N}$, $x \in c(S_i), y \in c(S_i)$.

It is equivalent to assuming that the set of chosen alternatives $c(S_i)$ is the set of *all* the best alternatives in S_i . Arrow (1959) and Richter (1966) have used strict revealed preferences, for instance.

Definition 1.6 (Weak Revealed Preferences). With weak revealed preferences, we only assume that chosen alternatives are not worse than unchosen alternatives.

- xR^0y if and only if there exists S_i , $x \in c(S_i)$, $y \in S_i$.
- xP^0y if and only if xR^0y and not yR^0x .
- xI^0y if and only if xR^0y and yR^0x .

It assumes that some unchosen alternatives might be among the best alternatives in S . Amartya Sen (1971) provides the link between strict and weak revealed preferences. Weak revealed preferences are particularly meaningful with choice functions, as the decision makers are forced to choose one alternative and therefore cannot reveal *all* their best alternatives if they have more than one, say if they are indifference. A simple example illustrates the differences between weak and strict revealed preferences. Take a grand set $X = \{x, y, z\}$, and choices observed in all subsets as $c(\{x, y, z\}) = \{x\}$, $c(\{x, y\}) = \{y\}$, $c(\{y, z\}) = \{y\}$, and $c(\{x, z\}) = \{x\}$. With strict revealed preferences, we deduce from the first choice xP^0y and xP^0z , from the second, yP^0x , from the third, yP^0z and from the fourth, xP^0z , which yield to xPy , yPx , xPz , and yPz . With weak revealed preferences, we deduce from the first choice xR^0y and xR^0z , from the second, yR^0x , from the third, yR^0z and from the fourth, xR^0z , which yield to xIy , xPz , and yPz . Weak revealed preferences yield a consistent revealed preference, whereas strict revealed preferences yield a logical problem if we interpret the choices in terms of welfare, as we have both that x is strictly revealed preferred to y and y is strictly revealed preferred to x . The choice of using weak or strict revealed preferences should depend on the context, and what interpretation of choice is the most sensible.

Definition 1.7 (Revealed Preferences on Consumption Data). With consumption data, we use the additional information given by prices and quantities to reveal preferences.

- $x_iR^0x_j$ if and only if $p_ix_i \geq p_jx_j$.
- $x_iP^0x_j$ if and only if $p_ix_i > p_jx_j$.
- $x_iI^0x_j$ if and only if $x_iR^0x_j$ and $x_jR^0x_i$.

This definition says that a bundle x_i is strictly preferred to a bundle x_j if acquiring x_j in situation i would have been strictly less costly. Two bundles are indifferent when they are chosen on the same budget line, where the budget line of situation i is $p_i x_i$. Hal R Varian (2006) summarizes revealed preferences on consumption data.

In all cases, $R^0 = P^0 \cup I^0$. Revealed preferences classify the different observed choices. They do not embed any sense of consistency or inconsistency. For that, we need a model of the preference of the decision maker. The standard assumptions on preferences yield *classical preference*.

Definition 1.8 (Classical Preference). A classical preference is a binary relation \succeq on X , which is reflexive, transitive, and complete.

- Reflexive: for any $x \in X$, $x \succeq x$;
- Transitive: for any $x, y, z \in X$, $x \succeq y$, and $y \succeq z$ implies $x \succeq z$;
- Complete: for any $x, y \in X$, $x \succeq y$ or $y \succeq x$. Completeness means that the binary relation ranks all the alternatives.

Classical preferences imply that in any choice set; there is a maximum (potentially non-unique however, or even infinite with consumption data). On finite data, having a classical preference is (trivially) equivalent to having a utility. On infinite data, Debreu (1954) have shown that classical preferences that are continuous are equivalent to having utility.³ That is, there exists a function u such that xRy if and only if $u(x) \geq u(y)$ and xPy if and only if $u(x) > u(y)$. The utility obtained is unique up to increasing positive transformations.

In practice, not all revealed preferences are classical preferences. Revealed preferences are classical preferences if and only if they are free of cycles (excluding cycles of indifference), a condition called *acyclicity*.

Definition 1.9 (Cycle). A cycle is a pair of alternative $x, y \in X$ such that xRy and yP^0x .

Definition 1.10 (Acyclicity of Revealed Preferences). Revealed preferences are acyclic if they contain no cycle of preferences. That is, if, for all $x, y \in X$, xRy implies not yP^0x .

If a family of \mathcal{N} situations generates acyclic revealed preferences, we say that the data is *rationalizable* by a classical preference. Besides, we can build a utility and may think about welfare. However, if the choices generate revealed preference cycles, the decision maker can no longer be modeled as a classical utility maximizer, which calls for a different notion of preference or a different approach to welfare analysis. Note that here, we have abused a bit the notion of a classical preference. Indeed, it might be the case that a revealed preference is acyclic, but that some comparisons have not been

³Continuity is defined in Debreu (1954) by let X be a completely ordered subset of a finite Euclidean space. For every $x \in X$, the sets $\{y \in X | y \leq x\}$ and $\{y \in X | y \geq x\}$ are closed. We are clearly in this case with consumption data endowed with the usual topology on \mathbb{R}^m .

observed, and so no revealed preference has been built, which means that completeness is violated. Here acyclicity only guarantees that it is possible to build a classical preference, not on X , but X restricted to alternatives that are compared in the data. With this restriction, the revealed preference built is unique. Without, there might be multiple ways to complete unobserved comparisons that are compatible with the observed ones. One condition that guarantees, in addition to acyclicity, that the revealed preference is complete is *full observability* (for finite data).

Definition 1.11 (Full Observability). Full observability states that we observe a choice for each choice set. That is, for all $S \in \mathcal{P}(X)$, there is a situation $i \in \mathcal{N}$, such that $S_i = S$.

Full observability is essential because some unobserved choices might yield to violations of the acyclicity conditions, even if observations on binary choices alone do not yield to observable violations. Indeed, we could, in theory, reveal a classical preference only with choices on choices set made of two alternatives. It is also useful to go beyond classical preferences and identify models of intransitive indifference and menu-dependence. De Clippel and Rozen (2018) explore in details why having all possible observations is necessary for assessing some models and why the conditions change when it is not the case. Full observability is only meaningful for finite data. By definition, with infinite data, full observability cannot be attained, and all conditions must be falsifiable on a finite subsample of all possible observed choices.

It is useful to concentrate on what particular case when we assume strict revealed preferences and full observability. The condition for a choice correspondence, and by extension, a choice function, to be rationalized by a classical preference is the *Weak Axiom of Revealed Preferences* (WARP hereafter), as shown by Arrow (1959).

Axiom 1.1 (Weak Axiom of Revealed Preferences (WARP)). For any $S \in \mathcal{P}(X)$ and $y \in S$, if there exists an $x \in c(S)$ such that $y \in c(T)$ for some $T \in \mathcal{P}(X)$ with $x \in T$, then $y \in c(S)$.

WARP says that if x is chosen when y is available, then certainly if y is chosen and x is present, x must be chosen too. It is equivalent to acyclicity with full observability. One way to understand WARP is to decompose it in two other axioms, α , and β , as shown by A. Sen (1969).

Axiom 1.2 ((ref:alpha)). For any pair of sets S and T in $\mathcal{P}(X)$ and any alternative x in S , if x is in $c(T)$ and $S \subseteq T$, then x must be in $c(S)$.

Axiom 1.3 ((ref:beta)). (β). For all pairs of sets S and T in $\mathcal{P}(X)$ and all pairs of elements x and y in $c(S)$, if $S \subseteq T$, then $x \in c(T)$ if and only if $y \in c(T)$.

Axioms α and β together are equivalent to WARP under full observability. α says that any alternative chosen in a large set should be chosen in any subset. β says that if two alternatives are chosen together in a set, they should be chosen together in any superset. Condition β is void on choice functions, which implies that WARP and axiom α are equivalent with choice functions.

For consumption data, the condition is again the acyclicity condition given in Definition 1.10. The condition is known as the *General Axiom of Revealed Preferences* (GARP) and has been proven by Hal R. Varian (1982).

Without full observability, the acyclicity condition on finite data is called *Strong Axiom of Revealed Preferences* (SARP). Revealed preferences often fail in practice, as Koo (1963) pointed out in an early study:

In an empirical study, it is not likely that one will find many individuals who are either entirely consistent or inconsistent.

It has proven remarkably prescient, and yield us now to study extensions of the classical paradigm.

1.2 Beyond Classical Preferences

One possibility to explore the failures of revealed preferences is to change the kind of preference it is mapped into. That is, to replace classical preferences with other kinds of preferences. A starting point is to relax the assumption of classical preferences: reflexivity, transitivity, and completeness.

Reflexivity does not have many implications on revealed preferences.

Transitivity of strict preferences has strong normative backing for welfare analysis, as the lack of transitivity means that it is potentially impossible to determine the best alternatives in a set, and thus makes it impossible to think about individual welfare. Violating the transitivity of strict preferences would also imply weird logical conclusions. It is possible, however, to relax transitivity of the indifference, using models of just-noticeable differences introduced by Luce (1956) and Fishburn (1970). Armstrong (1939) provides an early critique of the normative appeal of transitivity of the indifference. Revealed preference conditions for intransitive indifference models have been given by Schwartz (1976) and Aleskerov, Bouyssou, and Monjardet (2007), among others. Example 1.1 carries the idea behind intransitive indifference well.

Example 1.1 (Intransitive Indifference). Luce (1956) provides a famous example of why transitivity of the indifference might not be desirable, involving coffee and sugar:

Find a subject who prefers a cup of coffee with one cube of sugar to one with five cubes (this should not be difficult). Now prepare 401 cups of coffee with $\left(1 + \frac{i}{100}\right)x$ grams of sugar, $i = 0, 1, \dots, 400$, where x is the weight of one cube of sugar. It is evident that he will be indifferent between cup i and cup $i + 1$, for any i , but he is not indifferent between $i = 0$ and $i = 400$.

In short, decision makers might not perceive the difference between two alternatives under a certain threshold.

Completeness has been criticized quite early on from a normative standpoint too (see Aumann (1962), Bewley (2002)). Relaxing completeness does not prevent the finding of the best alternatives in a set of alternatives. Eliaz and Efe A. Ok (2006) and Aleskerov, Bouyssou, and Monjardet (2007) provide revealed preference conditions for incomplete preferences. Various reasons have been put forward in the literature to explain the emergence of incomplete preferences, which might be summarized in two. First, the decision maker might lack information on the alternatives available. Second, he might lack information on his preference over very rare alternatives.

The conditions for exploring models relaxing completeness or transitivity of the indifference are only meaningful on choice correspondences. Aleskerov, Bouyssou, and Monjardet (2007) have shown that on choice functions, the conditions are precisely the acyclicity condition of classical preferences. Thus, incomplete preferences and just-noticeable preferences require choice correspondence to be explored. To the best of our knowledge, only Costa-Gomes, Cueva, and Gerasimou (2016) have explicitly elicited a choice correspondence, albeit with a different modeling standpoint.⁴ From an empirical standpoint and with our modeling of choice correspondences, models of intransitive indifference with no transitivity conditions imposed on the indifference part of the preference are equivalent to models of incomplete preferences which impose transitivity on the strict part of the preference. For this reason, we will mainly talk about intransitive indifference.

It is possible to go even further with choice correspondence, and study models where the choice depends on the choice set, but not the preference. In these models, decision makers only approximately maximize their choice, and this approximation is choice set-dependent. Set-independence maximization has been criticized from a positive standpoint by Amartya Sen (1997). He argues that external conditions might influence the choice but not the underlying preference, and in particular moral considerations. Set-dependent models as studied here have strong link with models of intransitive indifference.

The intuition of the set-dependent models we will study here is as follows. Decision makers have difficulties in distinguishing alternatives that are close to each other and thus might choose alternatives that are not the best but close from the best. The threshold to distinguish depends on the set considered in these models, whereas it does not in just-noticeable difference models. It means that the other available alternatives influence the threshold. In Frick (2016), this threshold increases with set inclusion. She called it a *monotone threshold model*.⁵ Aleskerov, Bouyssou, and Monjardet (2007) introduced models that are relaxations of the model of Frick (2016). First, the threshold de-

⁴In this dissertation, we sometimes refer to Costa-Gomes, Cueva, and Gerasimou (2019) and Costa-Gomes, Cueva, and Gerasimou (2016). The latter is an earlier working paper version of the former, which contains more experimental results related to this dissertation. We try to cite the version the best related to the point made each time.

⁵Tyson (2018) has introduced models of set dependent choices that are strengthening of the monotone threshold model.

depends on the set considered, but there is no monotonicity condition imposed, it is *menu-dependent threshold model*. Second, the threshold depends on the set and the alternatives considered, is is a *context-dependent threshold model*.⁶ Aleskerov, Bouyssou, and Monjardet (2007) have shown that all the threshold models are not distinguishable from classical preferences with choice functions.

1.2.1 Relaxing Transitivity and Menu-Independence

Intransitive indifference models take their roots in the Weber-Fechner law of psychophysics that states that the threshold above which a difference between two stimuli is perceived is proportional to the original stimuli. It captures the idea that the magnitude of the difference between two measurable objects must be large enough to be noticed, as shown in Example 1.1.

The intuition behind just-noticeable difference shows why the transitivity of indifference is not a very compelling normative assumption. Fishburn (1970) is a survey of the theoretical literature on intransitive indifference models. Three models will be tested here, the original semi-order model of Luce (1956), the interval order model of Fishburn (1970) and the partial order model. Aleskerov, Bouyssou, and Monjardet (2007)'s Chapter 3 provide testable conditions for rationalizability by a semi-order, an interval order, and a partial order. The summary needed in this dissertation is given below. Dzielwski (2016) provides conditions to investigate just-noticeable difference on consumption data, but we will not dwell on that.

To the best of our knowledge, very few investigations on intransitive indifference exist in economics. It is despite Luce (1956) stating that there is a large body of literature in psycho-physics backing the notion of just-noticeable difference. Sautua (2017) rules out intransitive indifference as explaining his observations, but it is not a test of intransitive indifference models *per se*. This literature aims at eliciting the threshold that yields us to decide whether a stimulus is higher than another. For instance, which object is heavier between objects A and B. The common finding is the frequency of each object being perceived as the heaviest. The shift in perception of which object is the heaviest is not abrupt at the point where their weights are equal. Errors occur around this threshold, which disappears as the difference in weight becomes greater. They are two main difference with the investigation in economics. Preferences are subjective, rather than objective, in general. We usually observe one choice, and not several, between two alternatives, which does not allow us to build a frequency of choice. In general, carrying an investigation about the *unknown* preferences using a similar method does not seem realistic.

We introduce some definitions in order to characterize the different intransitive indifference models. In general, they differ on the restrictions they put on the indifference part of the relation. We go from the least structured to the most structured.

⁶These models are far from being compatible with every data, as shown in Appendix C.2.

Definition 1.12 (Quasi-Transitivity). A preference relation \succeq is *quasi-transitive* if the strict part of the preference relation is transitive, that is, \succ is transitive.

Definition 1.13 (Strong Intervality (aka Ferrers property)). A binary relation \succeq satisfies *strong intervality* if and only if x is strictly better than y and z is strictly better than t , implies that either x is strictly better than t or z is strictly better than y . Formally:

$$\text{for all } x, y, z, t \in X, (x \succ y \text{ and } z \succ t) \Rightarrow x \succ t \text{ or } z \succ y$$

Definition 1.14 (Semi-transitivity). A binary relation \succeq satisfies semi-transitivity if and only if x is strictly better than y and y is strictly better than z , implies that either x is strictly better than t or t is strictly better than z . Formally:

$$\text{for all } x, y, z, t \in X, x \succ y \text{ and } y \succ z \Rightarrow x \succ t \text{ or } t \succ z$$

Definition 1.15 (Intransitive Indifference). These conditions define three intransitive indifference models:

1. A *partial order* is a binary relation which is reflexive, asymmetric, and transitive. In our settings, it means that the revealed preference is reflexive and quasi-transitive.
2. An *interval order* is a partial order that satisfies the strong intervality condition. In our settings, it means that the revealed preference is reflexive, quasi-transitive, and the strict part P satisfies the strong intervality condition.
3. A *semi-order* is an interval order which satisfies the semi-transitivity condition. In our settings, it means that the revealed preference is reflexive, quasi-transitive, and the strict part P satisfies the strong intervality condition and the semi-transitivity condition.

In addition to relaxing transitivity of the indifference, we also relax menu-independence. In order to understand the relaxation of menu-independence we use, it is useful to introduce the link between the choice and the utility in these models.

Definition 1.16 (General Threshold Representation). A choice correspondence c on X admits a threshold representation if there exist two functions $u : X \rightarrow \mathbb{R}$ and $t : X \times X \times \mathcal{P}(X) \rightarrow \mathbb{R}^+$ such that for every S ,

$$c(S) = \{x \in S \mid \text{for all } y \in S, u(x) \geq u(y) + t(x, y, S)\}$$

u is the fully rational benchmark, i.e., represents a classical preference, and t the departure threshold of the representation, t is always positive valued. It depends only on combinations of x , y , and S , but one could imagine other dependency structures if more information about the context of choice is

available. This dependency structure captures interval orders and semi-orders, as well as attraction, decoy or choice overload effects from behavioral economics.⁷

In menu-independent just-noticeable difference models, the threshold between two alternatives may only depend on the alternatives. In menu-dependent models, it also depends on the set. The threshold function can depend on various combinations of x , y , and S , such as $t(x, S)$, for instance. When the threshold also depends on the sets, Aleskerov, Bouyssou, and Monjardet (2007) showed that these threshold representations reduce to menu-dependent ($t(S)$) or context-dependent ($t(x, y, S)$) models.⁸ Partial orders are the exception here, as far as we know, it is only representable with a multi-utility representation, as in Efe A. Ok (2002).

1.2.1.1 Partial Order

Eliaz and Efe A. Ok (2006), following Schwartz (1976) uses a weakening of classical revealed preferences, where now the chosen alternatives are *not worse* than the unchosen alternatives. This assumption on revealed preferences does *not* assume completeness of preferences anymore, or equivalently, does not assume transitivity of the indifference. This yields a consistency requirement called the Weak Axiom of Revealed Non-Inferiority (WARNI hereafter), which is a weakening of WARP.

Axiom 1.4 (Weak Axiom of Revealed Non Inferiority (WARNI)). For a given alternative y in S , if for all the chosen alternatives in S , there exists a set T where x is in T and y is chosen in T , then y must be chosen in S . This property should be true for all S and y .

for all $S \in \mathcal{P}(X), y \in S$, if for all $x \in c(S)$, there exists a $T \in \mathcal{P}(X), y \in c(T), x \in T \Rightarrow y \in c(S)$

WARNI states that if an alternative is not worse than all chosen alternatives, it must be chosen too. In other words, an alternative that is not chosen must be worse than at least one chosen alternative. A choice correspondence satisfies WARNI if and only if it is rationalized by a reflexive and quasi-transitive preference relation – i.e., a partial order. This preference \succeq is unique, and we have that $c(S) = \{x \in S \mid \text{there is no } y \in S, S \succ y\}$. Aleskerov, Bouyssou, and Monjardet (2007) provide an equivalent axiomatization of rationalizability by a partial order.⁹ Their axiomatization

⁷The attraction and decoy effects are the facts that introducing a third dominated alternative in the choice between two alternatives will change the relative probability of each alternative being chosen. See Landry and Webb (2017) for a general model of attraction and decoy effects. The choice overload effect is the fact that adding alternatives to the choice sets might yield worse welfare. See Chernev, Böckenholt, and Goodman (2015) for a review of the origins and effects of choice overload.

⁸More precisely, a threshold model with a threshold of the form $t(y, S)$ can be equivalently represented by a threshold model with a threshold of the form $t(x, y, S)$, as shown in Aleskerov, Bouyssou, and Monjardet (2007)'s theorem 5.1. We have kept the latter representation. A threshold model with a threshold of the form $t(x, S)$ can be equivalently represented by a threshold model with a threshold of the form $t(S)$, as shown in Aleskerov, Bouyssou, and Monjardet (2007)'s theorem 5.2. We have kept the latter representation.

⁹We do not formally prove this equivalence. Theorem 2 of Eliaz and Efe A. Ok (2006) restricted to the partial order shows that if WARNI is satisfied, it is possible to find a partial order that rationalizes the choice correspondence.

is linked to Amartya Sen (1971)'s decomposition of WARP in axiom α and β , rather than a direct weakening of WARP. Eliaz and Efe A. Ok (2006) have shown that a choice correspondence that satisfies WARNI also satisfies axiom α .

1.2.1.2 Interval Order

On a finite set of alternatives, interval orders can be represented according to a general threshold function where the threshold depends on one alternative (which one does not matter). The testable condition for interval order is *functional asymmetry*. More precisely, a choice correspondence is rationalizable by an interval order if and only if it is rationalizable by a partial order – i.e., it satisfies WARNI – and it satisfies the Functional Asymmetry axiom.

Axiom 1.5 (Functional Asymmetry (FA)). A choice correspondence satisfies *Functional Asymmetry* if some chosen alternatives in S are not chosen in S' , it must be that all chosen alternatives in S' that are in S are chosen in S' :

$$\text{for all } S, S' \in \mathcal{P}(X), c(S) \cap (S' \setminus c(S')) \neq \emptyset \Rightarrow c(S') \cap (S \setminus c(S)) = \emptyset$$

Interval orders are a strengthening of partial order and weakening of classical preferences. Again, the preference obtained is unique.

1.2.1.3 Semi-Order

On a finite set of alternatives, semi-orders can be represented according to a general threshold function where the threshold is constant. A choice correspondence is rationalizable by a semi-order if and only if it is rationalizable by an interval order and it satisfies the Jamison-Lau-Fishburn axiom.

Axiom 1.6 (Jamison-Lau-Fishburn (JLF)). A choice correspondence satisfies the *Jamison-Lau-Fishburn* axiom if S is made of unchosen alternatives in S' and S'' counts some chosen alternatives in S' , then chosen alternatives in S'' must be chosen in S (if they belong to it).

$$\text{for all } S, S', S'' \in \mathcal{P}(X), S \subseteq (S' \setminus c(S')), c(S') \cap S'' \neq \emptyset \Rightarrow c(S'') \cap (S \setminus c(S)) = \emptyset$$

Semi-orders are strengthening of partial orders and interval orders and a weakening of classical preferences. Again, the preference obtained is unique.

1.2.1.4 Monotone Threshold

In the monotone threshold model, the threshold depends only on S and is *non-decreasing* with set inclusion, i.e., $t(S') \leq t(S)$ whenever $S' \subseteq S$. When the size of the choice set is larger, the choice is less precise, and the threshold is larger, which is a simple way to take into account choice overload. Frick (2016) shows that *occasional optimality* characterizes the *monotone threshold* on choice correspondences.

Axiom 1.7 (Occasional Optimality). A choice correspondence satisfies occasional optimality if, for all $S \in \mathcal{P}(X)$, there exists $x \in c(S)$ such that for any S' containing x :

1. If $c(S') \cap S \neq \emptyset$, then $x \in c(S')$;
2. If y is in S , then $c(S') \subseteq c(S' \cup \{y\})$.

WARP requires that any alternative the decision maker chooses from S is optimal. Occasional optimality requires that at least some of the decision maker's choices from S be optimal.

1.2.1.5 Menu-Dependent Threshold

A natural weakening of the monotone threshold model is to allow for a non-monotone threshold. The threshold t depends on S , without constraints. Aleskerov, Bouyssou, and Monjardet (2007) provide the testable condition for a menu-dependent threshold model to rationalize a choice correspondence.

Definition 1.17 (Strict Cycle of Observation). A *strict cycle of observation* are n sets S_1, S_2, \dots, S_n in $\mathcal{P}(X)$ such that:

$$\begin{aligned} (S_1 \setminus c(S_1)) \cap c(S_2) &\neq \emptyset \\ (S_2 \setminus c(S_2)) \cap c(S_3) &\neq \emptyset \\ &\dots \\ (S_n \setminus c(S_n)) \cap c(S_1) &\neq \emptyset \end{aligned}$$

A strict cycle of observations is a cycle of strict revealed preferences.

Axiom 1.8 (Functional Acyclicity). A choice correspondence c satisfies *functional acyclicity* if it does not contain strict cycles of observations.

In revealed preferences terminology, functional acyclicity states that there are no cycles of strict revealed preferences. Aleskerov, Bouyssou, and Monjardet (2007) also provide the corresponding revealed preference: it is the *strict revealed preferences*. For subjects which satisfy functional acyclicity, we only directly elicit strict preferences, not indifference. Strict Preferences are acyclic

and therefore can be augmented using the transitive closure, but no condition is imposed on indifference. The strict preference is unique, but depending on whether completeness is imposed or not, the indifference part may not be.

1.2.1.6 Context-Dependent Threshold

In the context-dependent threshold model of Aleskerov, Bouyssou, and Monjardet (2007), the threshold depends on the menu S and on the alternatives x and y that are compared. They give the condition on choice correspondences for context-dependent rationalizability.

Axiom 1.9 (Fixed Point). A choice correspondence c satisfies *fixed point* if, for any $S \in \mathcal{P}(X)$, there exists an alternative x in S such that x in S' implies that x in $c(S')$ for any $S' \subseteq S$.

There is a link between fixed point and the α axiom (Axiom 1.2). α requires that all alternatives chosen in a set are chosen in any subset. Fixed point only requires that one alternative chosen in a set is chosen in any subset.

Aleskerov, Bouyssou, and Monjardet (2007) do not provide the corresponding preference, but it is easy to build it when fixed point and full observability are satisfied. Take the set of fixed points in the whole set X ($FP(X)$). It is made of the most preferred alternatives in X . All alternatives in $FP(X)$ are indifferent, i.e., they are always chosen together when they are both available. Now take the set of fixed points in $X \setminus FP(X)$, it is the most preferred alternatives in $X \setminus FP(X)$, and so on until the set of alternatives that are not fixed point is empty or a singleton. By definition, this procedure will finish, as every nonempty subset of X has a fixed point. It also implies that the constructed preference is unique and complete: it is a classical preference. Compared to strict revealed preferences obtained with functional acyclicity, this revelation of preferences also reveals indifference.

The three models introduced here are ranked, as their threshold representation clearly shows. A choice correspondence that satisfies occasional optimality satisfies functional acyclicity. A choice correspondence that satisfies functional acyclicity satisfies fixed point. Fixed point has a clear advantage: it provides a simple way to reveal the preference, despite a seemingly complicated threshold representation.

1.2.2 Robust Revealed Preferences

Another way to think about the failures of revealed preferences is to change the modeling of decision making. Decision makers use heuristics to make most of their decisions; they do not maximize a utility. Is it possible to use this knowledge about the decision processes to think about individual welfare? Tversky and Kahneman (1981) introduced a tool that is now widely used to think about

heuristics, *frames*, which roughly collects the conditions under which a choice is made. Indeed, a single decision maker might exhibit many heuristics at the same time (decoy and attraction effects, etc.), so modeling all of them in the decision making seems both complicated and of little benefit.

The best definition of a frame for our purpose comes from Rubinstein and Salant (2012):

A frame is a description or details that influence choice behavior, though it is clear to an observer that they do not affect the individual's welfare.

A famous example of a frame is the Asian disease problem as introduced by Tversky and Kahneman (1981).

Example 1.2 (Asian disease). Subjects in an experiment are required to choose one program from problem 1 and one program from problem 2. The wording is the same for both problems. The only differences lie in the description of the programs.

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

1. In problem 1: If Program A is adopted, 200 people will be saved. If Program B is adopted, there is $1/3$ probability that 600 people will be saved, and $2/3$ probability that no people will be saved. Which of the two programs would you favor?
2. In problem 2: If Program C is adopted 400 people will die. If Program D is adopted there is $1/3$ probability that nobody will die, and $2/3$ probability that 600 people will die. Which of the two programs would you favor?

In problem 1, program A is chosen by most respondents, whereas in problem 2, program D is adopted by most respondents, despite programs A and C and programs B and D sharing exactly the same consequences.

To think about welfare in behavioral economics, Salant and Rubinstein (2008) and Bernheim and Rangel (2009) both introduced frames in the modeling of decision makers.¹⁰ They did so to provide a unified tool to study welfare, as most of the literature in behavioral economics relied on *ad hoc* models of the heuristics. They assume that the frame does not have an impact on welfare. Caplin and Martin (2012), Rubinstein and Salant (2012), Benkert and Netzer (2018) also provide ways to use frames when assessing welfare.

The difference between the approaches of Salant and Rubinstein (2008) and Bernheim and Rangel (2009) is their stance on the use of the knowledge of the biases. Salant and Rubinstein (2008),

¹⁰A review of enhanced data sets, which include richer information than just final choices, is provided by Caplin (2016).

Rubinstein and Salant (2012), and Manzini and Mariotti (2014) argue for its use to correct the biases with a model of decision making. Bernheim and Rangel (2009), Bernheim (2009), on the other hand, promote the use “model-free” approaches, where the knowledge of the frame should only be used to consider the choice as valid, i.e., relevant for welfare analysis, or not, i.e., an error. This second approach is close in spirit with revealed preferences. The choice is still kept as an indicator of welfare. Pesendorfer and Gul (2009) provide a critique of the latter approach, which in essence points out that the modeler is almighty in his selection of the welfare-relevant choices. There is more, however, to the latter approach than the use of frames. For these reasons, in all the dissertation, all observed choices will be used. The frames will not be considered to drop any observed choice.

1.2.2.1 The Strict Unambiguous Choice Relation

Bernheim and Rangel (2009) define the following relation:

x is strictly unambiguously preferred to y (denoted xP^*y) if whenever x and y are both available (in some welfare-relevant frame), y is never chosen.

A fundamental difference between the strict unambiguous choice relation (SUCR hereafter) and the standard revealed preference relation is that with SUCR, multiple observations are considered jointly to reveal the preference, whereas, with revealed preferences, each observation is taken independently to do so.

Bernheim and Rangel (2009)’s Assumption 1 is full-observability. Whenever full-observability is satisfied, Bernheim and Rangel (2009)’s Theorem 1 shows that SUCR is acyclic, which is a useful property for welfare analysis. Under full-observability, because all binary choice sets are observed, a necessary condition for xP^*y is:

For some choice set (and for some welfare-relevant frame) where both x and y are available, x is chosen.¹¹

As a consequence, with full observability xP^*y if and only if x is strictly revealed preferred to y and y is never revealed preferred to x . With strict revealed preferences in Definition 1.5, we have:

Definition 1.18 (Strict Unambiguous Choice Relation (SUCR)).

$$xP^*y \text{ if and only if } xP^0y \text{ and not } yR^0x \quad (1.1)$$

¹¹This condition appears in Manzini and Mariotti (2014).

SUCR is, in essence, a robust version of revealed preferences. From Equation (1.1), it is evident that xP^*y implies that xP^0y . That is, the strict unambiguous choice relation is a subset of the strictly directly revealed preference relation, with strict revealed preferences.

1.2.2.2 The Transitive Core

A second model-free approach to welfare analysis is the *transitive core* (TC hereafter). Unlike SUCR, the transitive core of Nishimura (2018) is generated from another relation R^* (in practice, the revealed preference relation).

Definition 1.19 (Transitive Core (TC)). x is preferred to y (denoted $xTC(R^*)y$) if for all other options z , zR^*x implies zR^*y and yR^*z implies xR^*z .

Nishimura (2018) shows that TC makes recommendations that do not rely on arbitrary decisions from a modeler, and thus answers the critique of SUCR made by Pesendorfer and Gul (2009) and Manzini and Mariotti (2014). This conservative approach is in the same spirit as Bernheim and Rangel (2009), but there are differences between SUCR and TC. Nishimura (2018) presents theoretical examples where SUCR is coarser than TC, specifically for models of time preferences with relative discounting and regret preferences.

1.2.2.3 Other Welfare Relations

Several other welfare relations have been proposed in the literature that imposes little *ad hoc* model structure. In a recent paper, Apesteguia and Ballester (2015) suggest a welfare relation based on a measure of rationality called the “swaps index”. They provide a behavioral foundation for their index by identifying the axioms that characterize it. The corresponding welfare relation is found by choosing the preference order that is closest to (empirically) observed choices. To assess the closeness of a preference order, they look at the number of alternatives that must be ignored in each choice set to match the preferences implied by choices to the candidate preference order. This approach uses choice set frequencies to overcome ambiguities, so is less conservative than SUCR in making welfare assessments.

An additional axiomatization of welfare inference was suggested by Chambers and Hayashi (2012). They introduce an individual welfare functional, which is a function from a choice distribution to a relation on alternatives, and they provide axioms to characterize the individual welfare functional. Like Apesteguia and Ballester (2015), this approach uses frequencies to overcome ambiguities, which enables them to generate a linear order. However, unlike Apesteguia and Ballester (2015), the frequencies they use are stochastic choice probabilities. We are not going to use frequencies of choice in this dissertation, so we will not treat these relations.

1.3 Organization of the Dissertation

This dissertation expands on the literature on revealed preferences in two directions. First, in Chapter 2, with Daniel Martin, we ran the first empirical assessment of the strict unambiguous choice relation of Bernheim and Rangel (2009) and the transitive core of Nishimura (2018). Second, Chapter 3 provides a method to elicit a choice correspondence by making choosing maximal alternatives a dominant strategy. It aims at realigning the theoretical literature, which assumes in general that decision makers can choose sets with the empirical literature which usually observe the choice of an alternative rather than a set of alternatives. Chapter 4 illustrates the method with an experiment. The experiment also suggests that one reason we observe inconsistent choice is precisely that decision makers are forced to choose single alternatives.

Chapter 2

Predictive Power in Behavioral Welfare Economics

With Daniel Martin

2.1 Introduction

It is normatively appealing to retain choice as the basis for welfare assessments. One choice-based solution is to find a model of choice procedures, decision-making errors, or behavioral biases that explains observed choices and to use that model to conduct welfare analysis.¹ An alternative choice-based solution is to generate a relation from choices without imposing much *ad hoc* model structure and to use that relation to conduct welfare analysis (e.g., Bernheim and Rangel (2009), Chambers and Hayashi (2012), Apesteguia and Ballester (2015), Nishimura (2018)).

Given the nature of this divide, theoretical debate has emerged as to how much model structure is necessary to provide precise welfare guidance from inconsistent choices (Bernheim and Rangel (2009), Rubinstein and Salant (2012), Manzini and Mariotti (2014), Bernheim (2016)). It has been argued that “model-free” behavioral welfare approaches that are conservative in how they resolve the normative ambiguities produced by choice inconsistencies will have little to say about welfare in practice. While there are other normative criteria for policymakers besides the precision of welfare guidance, if an approach has little to say about welfare, then other considerations are likely to be moot.

We offer empirical evidence for this theoretical debate by determining, for standard data sets from the lab and field, the precision of welfare guidance offered by two behavioral welfare relations: the strict unambiguous choice relation (SUCR henceforth) proposed by Bernheim and Rangel (2009)

¹There are many such examples from the decision theory and behavioral economics literature, including Caplin, Dean, and Martin (2011), Manzini and Mariotti (2012), Rubinstein and Salant (2012).

and the transitive core (TC henceforth) proposed by Nishimura (2018). Both SUCR and TC are “conservative” in the sense that they do not attempt to resolve all of the normative ambiguities produced by choice inconsistencies, so are likely to be incomplete in their welfare guidance when behavioral biases impact choice.

SUCR and TC aim to be acyclic by excluding revealed preferences relation elements that produce cycles. SUCR retains a relation between x and y when x is strictly unambiguously chosen over y (denoted as xP^*y) which holds if and only if y is never chosen when x and y are available.² Alternatively, TC retains a relation element between x and y if all pairwise comparisons with an option z are consistent with the ordering implied by that relation element.

We evaluate whether these behavioral welfare relations offer precise welfare guidance by determining their “predictive power”, which is their ability to make sharp predictions.³ When a theory does not offer unique predictions, predictive power indicates how loose or tight its predictions are. Because SUCR and TC can be incomplete, they do not always pin down what an agent would select from a set of options, so their predictive power is in question. As an example, imagine the choices of $\{x\}$ from $\{x, y\}$, $\{x\}$ from $\{x, y, z\}$, $\{x\}$ from $\{x, a\}$, and $\{a\}$ from $\{x, y, a\}$. From these choices, SUCR says that xP^*y , xP^*z , and aP^*y . For the choice set $\{x, y, z\}$, SUCR predicts that just x should be selected. On the other hand, for other choice sets, such as $\{x, a\}$, SUCR predicts that any option could be selected.

Predictive power is a useful way to evaluate the precision of welfare guidance because the predictions of a relation correspond to what is welfare optimal for that relation. For instance, if a welfare relation predicts that just one option could be selected from a choice set, then it has both maximal predictive power and offers the most precise welfare guidance. However, if a welfare relation predicts that any option could be selected from a choice set, then it has minimal predictive power and offers no welfare guidance. In the previous example, SUCR offers very precise welfare guidance from $\{x, y, z\}$ as the individual welfare optimum for that choice set is x , but it offers no welfare guidance from $\{x, a\}$.

One natural measure of predictive power is the number of options that are predicted to be chosen, where the highest possible predictive power corresponds to a value of 1 (a single option), and larger values represent less predictive power. However, we use Selten’s index (Selten (1991)) as our primary measure of predictive power instead because it has a theoretical grounding, has been used for related questions in the literature, and accounts for the number available alternatives. This index, axiomatized in Selten (1991), is designed explicitly for theories that predict a subset of possible outcomes. With Selten’s index, the proportion of choices that a theory predicts successfully within-sample is reduced by the “size of the area”, which is how many outcomes are consistent with

²Masatlioglu, Nakajima, and Ozbay (2012) provide an example of where SUCR and their model provide different welfare guidance, so SUCR is not free of model structure.

³For other applications of predictive power in empirical revealed preference analysis, see Manzini and Mariotti (2010), Beatty and Crawford (2011), Andreoni, Gillen, and Harbaugh (2013), Dean and Martin (2016), Boccardi (2018).

a theory. We calculate the size of the area by determining the fraction of options in a choice set that are predicted to be chosen.⁴

We test SUCR and TC’s predictive power for two types of data: from the lab, a set of choices from an incentivized experiment; and from the field, a set of scanned grocery purchases. The former is composed of choices from menus of payment plans for 102 students, which comes from an experiment carried out by Manzini and Mariotti (2010).⁵ The latter is composed of choices from budget sets for 1,190 single-person households over ten years, which comes from Nielsen’s National Consumer Panel (NCP) – formerly known as the Homescan Consumer Panel.⁶ The last data set requires us to make some assumption about the data, in particular, separability with respect to other expenditures, which reduces the “model-free” quality of SUCR and TC. These assumptions are standard in the empirical revealed preference literature, and we do not believe they hamper our analysis of the validity of SUCR and TC compared to RP, as we make the same assumptions for both.

We selected these data sets for four reasons. First, both are representative of widely used types of data in the economic literature. Second, in both data sets, individuals make inconsistent choices: for the experimental data, 53% of individuals make choices that generate revealed preference cycles, and for the consumption data, 100% of individuals exhibit revealed preference cycles.⁷ Third, both have unique features that make them rich enough to effectively test predictive power: the experimental data contains satisfies full observability, and the consumption data contains a large number of individuals and observations per individual. Fourth, they are quite different from each other in terms of individual demographic characteristics, choice settings, and alternatives.

For both data sets, we find that SUCR and TC have a high level of predictive power. In the experimental data, the average number of predicted options is 1.32 for SUCR and 1.38 for TC, and in the consumption data, it is 1.33 for SUCR and 1.65 for TC. In the experimental data, the average value of Selten’s index (which can range from 0 to 0.58 in this application) is 0.46 for SUCR and 0.44 for TC, and in the consumption data, the average value of Selten’s index (which can range from 0 to 0.96 here) is 0.95 for SUCR and 0.94 for TC.⁸

To learn when and why SUCR and TC have high predictive power, we study two properties of revealed preference (RP) that should correspond with their predictive power: the number of direct RP cycles and the fraction of all RP cycles that are direct. For the experimental data, we use strict revealed preferences as in Definition 1.5, and for the consumption data, we use revealed

⁴Because SUCR and TC always predict successfully within-sample, the value of Selten’s index in our application is determined entirely by the average size of the area.

⁵We are very grateful to the authors for providing this data to us.

⁶Data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

⁷We restrict our subsequent analysis to those subjects with cycles in their revealed preferences.

⁸Predictive power, as determined by Selten’s index, is higher in the consumption data in part because choice sets are larger on average.

preferences as in Definition 1.7.⁹ We say there is an RP cycle if there exists x_1, x_2, \dots, x_n such that $x_1 P x_2, \dots, x_n P x_1$ and that an RP cycle is “direct” if $x P y$ and $y P x$, that is, if it has a length of 2.¹⁰

It is worth noting that while cycle length is typically not considered in revealed preference analysis beyond distinguishing between violations of the Weak Axiom of Revealed Preference (WARP) and the Strong Axiom of Revealed Preference (SARP),¹¹ cycle length plays a critical role in this application.¹² For example, in determining the predictive power of SUCR and TC, the number of direct RP cycles is more important than the number of RP cycles of any other length or even the total number of RP cycles. The number of direct RP cycles is very informative about predictive power because SUCR and TC do not contain any relation elements that generate direct RP cycles, but can contain relation elements that are a part of longer length RP cycles.

That said, the number of longer length RP cycles does matter, but what matters most is their number in proportion to the number of direct RP cycles. If there are a lot of RP cycles of longer length relative to the number of direct RP cycles, then this is a problem for both SUCR and TC. For one, SUCR will be cyclical (P^* will contain cycles) if revealed preference cycles remain after removing all revealed preference relation elements that are in direct RP cycles, so SUCR is more likely to be cyclical if there are many RP cycles of longer length relative to the number of direct RP cycles. Besides, TC excludes additional revealed preference relation elements beyond those that are in direct RP cycles, so its predictive power is likely to be lower if there are many RP cycles of longer length relative to the number of direct RP cycles. Thus, for the predictive power of TC, the fraction of all RP cycles that are direct is likely to matter in addition to the number of direct RP cycles.

Because determining the number of RP cycles of longer length can become computationally burdensome, we generate a bound on the fraction of all RP cycles that are of length two by dividing the number of RP cycles of length two by the sum of all RP cycles of length two and three. We call this measure the “directness index”, and to the best of our knowledge, this measure is new to the literature.

As expected, we find that the *number* of direct RP cycles (RP cycles of length two) is highly and negatively correlated with Selten’s index for both SUCR and TC in our data sets. In the experimental data, the correlation for SUCR is -0.98, and for TC is -0.92, and in the consumption data, the correlation for SUCR is -0.75, and for TC is -0.81. We also find that the *fraction* of revealed preference cycles that are direct is highly and positively correlated with Selten’s index

⁹We use the strict RP relation P because we never observe more than one option selected from a choice set and never observe two selected bundles that could have been purchased at the identical expenditure. Additionally, we would not have cycles of length two with weak revealed preferences, a measure that is important to understand predictive power, as we will show.

¹⁰When counting cycles, we avoid double-counting by requiring x_1, x_2, \dots, x_n to be distinct and assuming that any re-ordering of x_1, x_2, \dots, x_n is the same cycle.

¹¹Direct RP cycles are violations of both WARP and SARP, whereas RP cycles of longer lengths are only violations of SARP.

¹²An exception is V. Aguiar and Serrano (2017), who consider the implications of cycles of different lengths relative to the Slutsky Matrix.

for TC. In the experimental data, the correlation for TC is 0.78, and in the consumption data, the correlation for TC is 0.69. These relationships are also strong, positive, and significant in regressions that control for the number of direct RP cycles. As measured by R^2 , variation in these two factors explains 89% of the variation in Selten's index for TC in the experimental data, and 67% in the consumption data.¹³

We find that the *fraction* of direct RP cycles is strongly and positively correlated with the acyclicity of SUCR. In the consumption data, 79% of individuals have acyclic SUCR, and the correlation is 0.38 between the directness index and a dummy variable that takes a value of 1 when SUCR is acyclic.

In addition to offering an answer to when and why SUCR and TC perform well in practice, both measures (the number of direct RP cycles and the directness index) are relatively quick to calculate. When approaching a new data set, it is easy to assess whether conservative model-free approaches to behavioral welfare economics are likely to offer precise welfare guidance.

To the best of our knowledge, this paper presents the first non-parametric empirical application of SUCR and the first empirical application of TC.¹⁴ We also introduce predictive power as a tool for evaluating behavioral welfare relations. Based on a standard measure of predictive power, we help to provide an answer to the question of how much model structure is necessary to provide precise welfare guidance. For the standard choice data sets we consider, it appears that one can give precise welfare guidance without imposing many assumptions – on the form of utility, on the nature of the behavioral biases, or on which choice sets to consider.

In Section 2.1.1, we briefly introduce talk about the empirical literature related to SUCR and TC. In Section 2.2, we describe the two data sets. In Section 2.3, we provide results for both data sets. We conclude with a brief discussion in Section 2.4.

2.1.1 Empirical Findings in Behavioral Welfare Economics

Bernheim, Fradkin, and Popov (2015) provide the first empirical implementation of SUCR to choice data.¹⁵ They study the impact of making one retirement savings option the default, and because individuals appear to make inconsistent choices as the default option changes, they use SUCR to identify the welfare impacts of such a change. However, to generate these welfare judgments, they make additional assumptions about the parametric form of utility and how different aspects of the choice correspondence relate to frames. We make no such additional assumptions, so our results are

¹³We do not have an *a priori* reason to believe that the directness index should matter for the predictive power of SUCR when controlling for direct RP cycles.

¹⁴As discussed in section 2.1.1, there are existing parametric empirical applications of SUCR.

¹⁵An application of concepts from Bernheim and Rangel (2009) also appears in Ambuehl, Bernheim, and Lusardi (2014).

better situated to address the question of whether precise welfare assessments can be made with a limited model structure.

Apestequia and Ballester (2015) present an empirical application of the swaps index as a measure of rationality, but they do not provide results on the corresponding welfare relation. One challenge in empirically assessing the swaps welfare relation is that it may not be uniquely identified for data sets that do not have full observability, unlike the relations suggested by Bernheim and Rangel (2009) and Nishimura (2018). However, Apestequia and Ballester (2015) formally prove that the mass of data sets for which the swaps welfare relation is not unique has mass zero, and when the welfare relation is not unique, the different welfare relations are likely to be very close to each other and coincide in the upper part of the rankings.

Finally, the results for our consumption data are not entirely unexpected, as Dean and Martin (2016) show for a panel of grocery store scanner data that households are “close” to being rational in the sense that the minimal cost to make a revealed preference relation acyclic is relatively small. However, there are three ways in which the high predictive power of SUCR and TC for our consumption data is surprising, even in light of their findings. First, Dean and Martin (2016) consider the minimal cost to make a revealed preference relation acyclic, whereas SUCR and TC remove all ambiguous comparisons, which is in general, much more conservative. Second, our panel is eight years longer than theirs, so it provides a much stricter testing ground as it contains five times more observations. Third, and most importantly, even if only a few revealed preference relation elements need to be removed from a relation to make it acyclic, there is no guarantee that such a relation will have high predictive power.

2.2 Data

We use two very different data sets for our non-parametric applications of Strict UCR. The first one comes from an experiment carried out by Manzini and Mariotti (2010) and consists of choices among different sequences of delayed payments. The second one comes from the Nielsen Consumer Panel (NCP) and consists of grocery purchases recorded by the marketing firm Nielsen for over ten years. Among the many differences between these data sets are the subject pools (students versus shoppers), the choice setting (lab versus field), and the choice alternatives (choices from menus versus choices from budgets).

Despite these differences, both are representative of widely used types of data in the economic literature. Data from experiments in which subjects are asked to choose among delayed payments appear in many papers because they can be helpful when studying time-inconsistencies and time preferences (see Frederick, Loewenstein, and O’Donoghue (2002)). Grocery store scanner data appears in several papers in the economics literature because it offers both price and quantity

Table 2.1: Installment plans

Delay	2 installments				3 installments			
	I2	D2	K2	J2	I3	D3	K3	J3
3 months	16	32	24	8	8	24	16	8
6 months	0	0	0	0	16	16	16	8
9 months	32	16	24	40	24	8	16	32
Total	48	48	48	48	48	48	48	48

information at the UPC level across a wide range of households living in different markets with varying demographic characteristics. For instance, M. Aguiar and Hurst (2007) use grocery store scanner data to study the purchasing habits of retirees.

2.2.1 Experimental Data

The task that subjects undertook in this experiment was a simple choice task: subjects were asked to pick their preferred payment plan from a list of options. All payment plans were sequences of installment payments that were delayed by 3, 6, or 9 months. In each choice that a subject made, all of the listed plans had either two or three installments. In general, there were four types of plans, which were called the increasing plan (I), the decreasing plan (D), the constant plan (K), and the jump plan (J). For all plans, the total payment was 48€. The exact payments and delays for both sets of options are presented in Table 2.1. At the end of the session, for each subject, one choice was drawn at random and implemented. The subjects all received their delayed payments at their due dates. Additional details are available in Manzini and Mariotti (2010).

A unique feature of the experiment of Manzini and Mariotti (2010) is full observability: subjects were asked to choose from all possible subsets of choice options, which can be interpreted as eliciting the entire choice function.¹⁶ Data with the property of full observability are appealing for two reasons. First, Strict UCR is guaranteed to be acyclic for such data. Second, such data provide a stringent test of the predictive power of Strict UCR.

Because subjects were asked to choose from all subsets for two sets of four plans, they made a total of 22 choices (each set of four plans corresponded to 11 choices). In the treatment where choices were incentivized, 102 individuals completed the experiment.¹⁷

¹⁶Presenting all possible subsets is combinatorially challenging. For n alternatives, the number of choice sets is $2^n - n - 1$. In Chapter 4, another experiment with full observability is presented. Another example of such an experiment is provided by Costa-Gomes, Cueva, and Gerasimou (2016). None of these examples goes above a universe of more than five alternatives.

¹⁷Choices were not incentivized for an additional 54 subjects, so we do not include them in our analysis. However, the results do not qualitatively change if we also include those subjects in our analysis.

2.2.2 Consumption Data

This data set is a balanced panel of purchases for single-person households that we have extracted from Nielsen's National Consumer Panel (NCP). NCP was formally known as the Homescan Consumer Panel because these grocery purchases are recorded using a scanner. There are a growing number of papers that analyze NCP data.¹⁸

A unique feature of NCP is the duration of the panel. For the single-person households we study, the data set contains information on grocery purchases over ten years. The length of this panel means that we have many observations, which allows us to perform a stringent test of the predictive power.

2.2.2.1 Analysis Sample: Panelists

To construct our analysis sample, we start with purchases made by 140,827 households during a ten-year window (from 2004 to 2013). The full data set contains records for purchases of 565,583,696 goods from 98,684,440 store trips, and the purchases correspond to 3,692,767 Universal Product Codes (UPCs).

From these observations, we extracted a balanced panel of 1,190 singles who satisfy the following criteria over the entire ten years:

1. Made purchases every month;
2. Stayed single;
3. Did not move to a different market area (as defined by Nielsen);
4. Did not retire.

While these restrictions may reduce the representativeness of our sample, the motivation for using such criteria is to keep preferences as stable as possible within each household over the ten years we study. The demographic characteristics of the samples are given in the next subsection, as well as an assessment of their representativeness. For instance, we look at singles who stayed single because Dean and Martin (2016) find that singles and married couples have different levels of choice inconsistency. Also, we look at singles who do not retire because M. Aguiar and Hurst (2007) find that retirement influences consumption patterns.

Nielsen registers purchases for a wide variety of products. To avoid products that can be stored for long periods, we have restricted ourselves to purchases of edible grocery products. This restriction reduces the original data to 365,014,702 goods purchased during 55,670,551 store trips and with

¹⁸As of June 2019, 134 working papers released by the Kilts Center use NCP. The current list of such papers can be found at <http://www.ssrn.com/link/Chicago-Booth-Kilts-Ctr-Nielsen-Data.html>.

Table 2.2: Average hours worked per week.

	<30 hours	30-35 hours	>35 hours	Not employed
Analysis sample over ten years	9.48%	3.52%	44.01%	42.99%
30+ year old in US (2004)	10.72%	4.81%	47.23%	37.23%

Source

Table 19 of the CPS Labor Force survey. http://www.bls.gov/cps/cps_aa2013.htm

1,436,818 different UPCs. By further restricting the data of our balanced panel to singles, we end up with 5,897,440 goods purchased during 1,317,467 store trips, accounting for 329,753 UPCs.

For the singles in our analysis sample, the average expenditure per month and per panelist on the goods we have kept is \$235.05, whereas the average total expenditure per month and panelist is \$427.27 for all households and goods in the NCP over these ten years.

2.2.2.2 Analysis Sample: Demographic Characteristics

All of the subjects who participated in the experiments of Manzini and Mariotti (2010) were Italian university students. On the other hand, the panelists in our consumption data are residents of the US, older, and mainly working full-time or close to full-time.

For the analysis sample of our consumption data, the median age in 2004 is 56 years, and the youngest panelist in 2004 is 30 years old. Among individuals in the US who were 30 years old and above in 2004, the median age is 50.¹⁹

As shown in table 2.2, a majority of individuals in our analysis sample are working, and a plurality works more than 35 hours per week. There is, however, a substantial fraction that is not employed (42.76% on average over the ten years), and this rate is higher than for individuals in the US who were 30 years old and above in 2004 (37.23%). The sample is skewed towards people already retired. While we have excluded individuals that experience a change from employment to retirement, we have not removed those who are retired or inactive throughout the ten years.

The median income of the analysis sample is between \$30,000 and \$35,000, which is slightly lower than the median income of individuals in the US who were 30 years old and above in 2004, as shown in Table 2.3. The level of education of our sample is slightly higher than this group, as Table 2.4 shows.

In the experiments of Manzini and Mariotti (2010), the subjects were a roughly even mix of men and women (see footnote 9 of Manzini and Mariotti (2010)). In the analysis sample of our consumption

¹⁹Data from the Current Population Survey (CPS) for 2004. <http://www.census.gov/population/age/data/2004comp.html>

Table 2.3: Income quartiles. The original data has income brackets, so the midpoint is used.

	25 th	50 th	75 th
Analysis sample over ten years	\$17,500	\$32,500	\$47,500
30+ year old in US (2004)	\$26,250	\$38,750	\$56,250

Source

Annual Social and Economic (ASEC) Supplement of the CPS. http://www2.census.gov/programs-surveys/cps/tables/pinc-03/2005/new03_010.txt

Table 2.4: Level of education. The degree considered is the highest received, so some individuals in the "no college" category might have been to college, but did not get their degree.

	College Degree	No College Degree
Analysis sample over ten years	46.55%	53.45%
30+ year old in US (2004)	43.25%	56.75%

Source

Annual Social and Economic (ASEC) Supplement of the CPS. http://www2.census.gov/programs-surveys/cps/tables/pinc-03/2005/new03_010.txt

data, 733 out of the 1,190 panelists are women, a proportion of 62%. In the US population, the fraction of women among individuals aged 30 and older was 52.34% in 2004.²⁰

2.2.2.3 Analysis Sample: Bundles

For a given month, each panelist has a corresponding bundle, made of six goods with quantities expressed in ounces. In order to construct bundles, we aggregate all purchases made during a month and aggregate the purchases into six categories given by Nielsen: alcoholic beverages, dairy products, deli foods, dry groceries,²¹ frozen food, and packaged meat. Average budget shares for these product categories are given in table 2.5. Aggregation over a month is done for two reasons: first, to compensate for the fact that panelists do not in general shop every day; and second, to assuage concerns about the storage of products. Because the units of measure are not necessarily the same between UPCs, we have first converted every product quantity into ounces (either fluid or solid), so that each aggregated good is quantified in ounces.

Building bundles by aggregating over categories and periods is standard in the literature that uses scanner data. For instance, Dean and Martin (2016) build similar bundles to perform a revealed preference analysis using scanner data; Hinnosaar (2016) aggregates beer into one homogeneous

²⁰US Census Bureau, CPS survey, Annual Social and Economic Supplement, 2004.

²¹The category dry grocery has a subcategory of pet food which we have removed. First, it is not edible, and second, there should be little substitution between pet food and human food.

Table 2.5: Average budget shares (expenditure on a product/good category in proportion to total expenditure) in a month.

Product	Average	Standard deviation ^a
Alcoholic beverages	5.71%	14.26%
Dairy products	16.25%	14.13%
Deli foods	2.56%	4.59%
Dry groceries	62.33%	19.89%
Frozen food	10.33%	10.15%
Packaged meat	2.82%	4.51%

^a Standard deviation in proportion from the average.

good, and Handbury (2013) study inflation with price indices built similarly.

2.2.2.4 Analysis Sample: Prices

The panelists are divided by Nielsen in 58 markets, which correspond roughly to large metropolitan areas of the United States. These markets and the number of panelists in each market are given in Figure 2.1. For each market, we have built a price vector, which is a unit price for each aggregated good expressed in dollars per ounce. To build this price vector, we use a “Stone” price index:

$$P_{Jt} = \sum_{i \in J} w_{it} p_{it}$$

where P_{Jt} is the price index for good category J in period t , w_{it} is the budget share for UPC i in period t , and p_{it} is the mean price for UPC i in period t .²²

We know that there is measurement error in prices, in particular, because panelists sometimes enter prices themselves. Indeed, Nielsen uses the following data collection methodology: each panelist has a scanner at home and scans all purchases once home. Nielsen matches a price to the UPC by linking these purchases to a database of store prices. If a price is missing, the panelist is required to input the price by hand. To incentivize the panelists to make correct entries, Nielsen has different cash reward programs, but some price entry errors are inevitable. To reduce the impact of these and other price measurement errors, we take two steps. First, we use purchases from the entire panel to construct market prices, not just purchases from our analysis sample. Second, we do not consider entries in the upper 2.5% and lower 2.5% of the price distribution for a product category in a period.

²²Dean and Martin (2016) do not find significant differences in revealed preference violations when using Stone, Laspeyres, or Paasche indices.

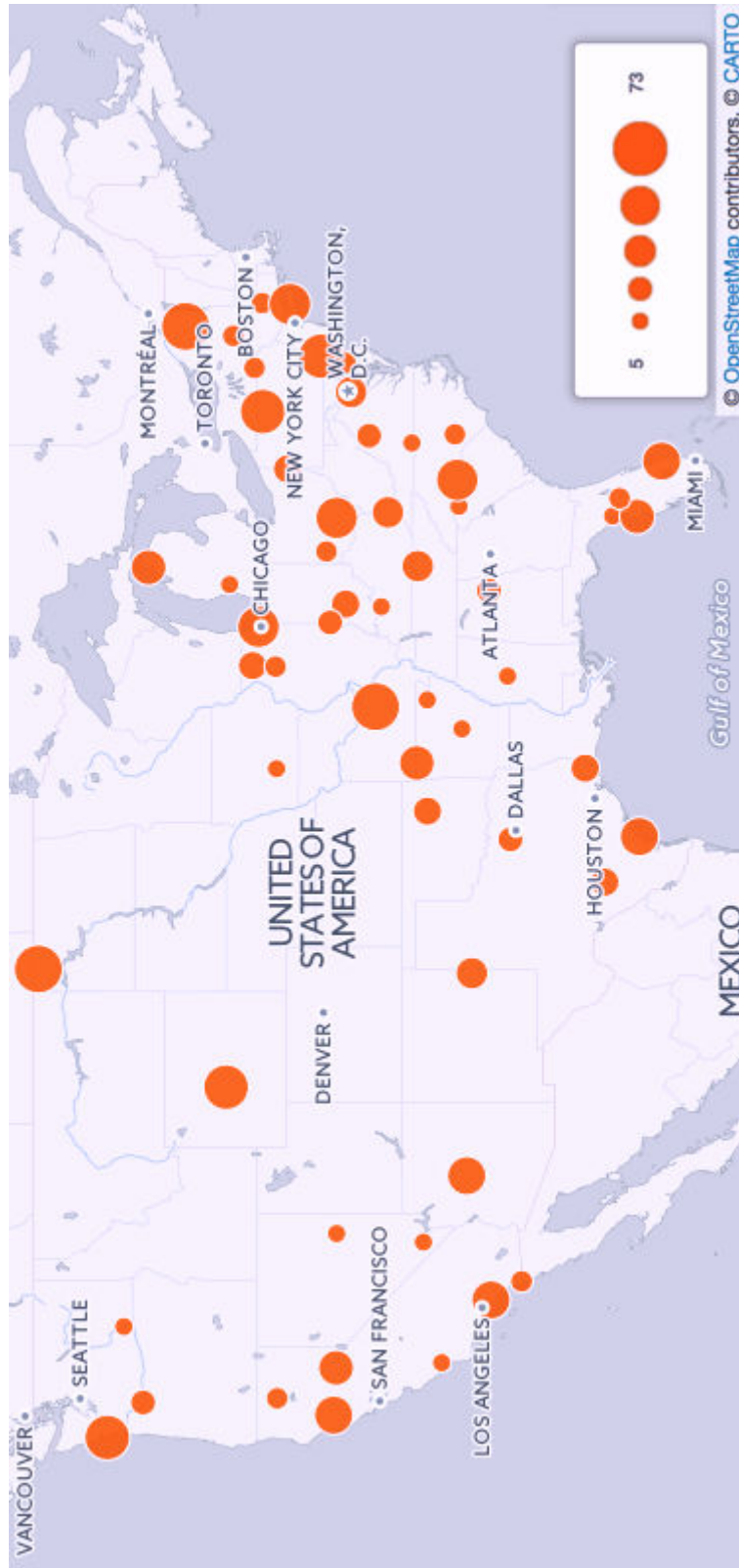


Figure 2.1: Individuals in the consumption data by market. The size of a bubble is proportional to the number of individuals in a given market.

2.2.2.5 Additional Considerations

Of course, grocery purchases are just one component of a household’s regular expenditures. An implicit assumption made when considering the consistency of these choices is separability between grocery purchases and the rest of a household’s expenditures. A justification for separability is that households may have a separate grocery budget. While strong, separability is a standard assumption in applications of revealed preference techniques to consumption data (for instance, see Koo (1963), Blundell, Browning, and Crawford (2003), Dean and Martin (2016)).²³ This assumption is backed by the findings about mental accounting (see Thaler (1985)). Individuals tend to silo their budget and expenditures on different kinds of goods.

Another standard assumption is that all panelists from the same market face the same prices in a given period. This assumption is necessary because if a household does not buy from a product category in a given period, prices are not identified for that category. Because we are using market prices, our analyses capture the impact of sustained and widespread price changes, not very temporary and local ones. Once again, this is a standard assumption in the applied revealed preference literature.

The last important assumption made for empirical testing is the stability of preferences over time, which is needed to make comparisons across periods. If preferences were to change, then having violations of revealed preferences would only mean that preferences have changed and would not be informative *per se*. While this assumption is also standard in the applied revealed preference literature, we recognize that it could potentially impact our results. However, even if preferences are indeed unstable over time, this should work against the precision of SUCR and TC, which would make the test of predictive power even tougher.

2.3 Results

In this section, we first determine the proportion of individuals who have choices that exhibit revealed preference cycles. For such individuals, we then determine the proportion that has cycles in SUCR and TC, the completeness and predictive power of these relations, and the properties of revealed preferences that are correlated with the predictive power of these relations.

2.3.1 Inconsistencies in Revealed Preferences

As discussed previously, a standard marker for choice inconsistency is the presence of cycles in the preferences revealed by choice. We say that x is (strictly) revealed preferred to y (denoted xPy) if

²³However, this does impose some model structure, which is another reason SUCR and TC are not entirely “model-free” in our application.

x is chosen when y is available and that there is an RP cycle if there exists x_1, x_2, \dots, x_n such that $x_1 P x_2, \dots, x_n P x_1$. If such a cycle exists, then the choices that generated it cannot be rationalized with a single, stable utility function, which is the standard tool for conducting welfare analysis.

For both of our data sets, a majority of individuals have choices that generate at least one RP cycle, as shown in Table 2.6. In the experimental data, 53% of individuals have RP cycles for at least one installment plan. In the consumption data, all 1,190 individuals have RP cycles.

The breadth of RP cycles we observe is consistent with findings in the empirical literature on revealed preference testing. In the laboratory experiments of Choi, Fisman, et al. (2007), around 35% of subjects have RP cycles for choices from allocations over risky assets, and in the large-scale field experiment of Choi, Kariv, et al. (2014), around 90% of subjects exhibit RP for a similar choice task.

For consumption data, there is a long history of papers that detect RP cycles. In one of the earliest computer-based studies of consumption data, Koo (1963) examined a panel of food purchases from 1958 for 215 Michigan households and concluded: “In an empirical study, it is not likely that one will find many individuals who are either entirely consistent or inconsistent.” This prediction has held for subsequent studies, including a paper by Dean and Martin (2016) which finds that around 71% of households exhibit RP cycles in a two-year balanced panel of grocery purchases.

Table 2.6: Percent of individuals that have cycles in RP, SUCR, and TC.

Data	Number of individuals	RP	SUCR	TC
Experimental	102	53%	0%	0%
Consumption	1,190	100%	21%	0%

2.3.2 Inconsistencies in SUCR and TC

SUCR and TC are designed to produce welfare guidance that is free of cycles. However, they are only guaranteed to be acyclic under certain conditions. As mentioned previously, a requirement that guarantees SUCR will be free of cycles is full observability. TC is free of cycles when choices from all binary choice sets are observed, which is a weaker condition. However, this condition is sufficiently strong to ensure that the underlying revealed preference relation will be complete.

The experimental design meets both conditions, so SUCR and TC are acyclic in the experimental data. On the other hand, neither condition is satisfied with the consumption data, so their acyclicity is in doubt in this data set. As shown in Table 2.6, SUCR has cycles for 21% of individuals in the consumption data. It represents a substantial reduction from the 100% of individuals who have RP cycles in that data set. TC achieves an even more substantial reduction: even though acyclicity is not guaranteed for TC, it never contains cycles in this data set.

Because SUCR and TC are only needed for welfare guidance when individuals have cyclic revealed preference relations, we only consider individuals that have RP cycles in the remaining analyses. It does not restrict our consumption data at all, but it means that we keep only 53% of experimental subjects, which leaves a total of 54 subjects in our analysis sample.

As a baseline, we keep those individuals in the consumption data who have SUCR cycles. However, because these cycles could potentially distort our assessment of the completeness and predictive power of SUCR, we also provide results without those individuals as a robustness check.

2.3.3 Completeness of SUCR and TC

A relation \succeq is “complete” if for all x and y in the grand set of alternatives X , either $x \succeq y$ or $y \succeq x$. One of the ways to measure the completeness of a relation is to measure the number of relation elements it contains, and this measure can be normalized by dividing it by the number of relations in a classical preference. A classical preference has $\frac{|X|(|X|-1)}{2}$ relation elements.²⁴

Because SUCR excludes all relation elements that are a part of direct RP cycles, Rubinstein and Salant (2012) and Manzini and Mariotti (2014) have argued that SUCR has the potential to be quite incomplete. However, on average, we find that SUCR and TC are far from incomplete in our data sets, as shown in Figures 2.2 and 2.3.

In the experimental data, because individuals make choices separately from two sets of four options, a classical preference would have 12 relation elements. For individuals with cyclic RP, the average number of relation elements for SUCR is 9.5, which is 80% of the comparisons in a classical preference. For TC, the corresponding figures are 9.1 and 76%.²⁵ There is heterogeneity in the extent of completeness: none of these subjects have complete SUCR and TC relations, but 52% are one relation element short with SUCR and 46% are with TC. An additional 13% are two relations short with SUCR and 9% with TC. Some individuals, however, have a half or less of a classical preference (9% of with SUCR and 13% with TC).

In the consumption data, a classical preference would have 7,140 elements. To determine this number, we take the grand set of alternatives X to be the set of all bundles that an individual has chosen at some point.²⁶ In theory, choices in the consumption data also generate revealed preference content about bundles that are never chosen, but such revealed preference content will always be included in SUCR and TC, so including it would inflate our assessment of the completeness of

²⁴If a relation contains direct cycles, then this number can be exceeded (up to twice this number). In our analysis, RP will often exceed this number, but SUCR and TC, which do not contain direct cycles, will never exceed it.

²⁵The difference between the average number of relations for SUCR and TC is significant (the two-sided paired t-test p-value is 0.0016). For a KS-test of equality of the distributions of SUCR and TC completeness, the p-value is <0.001.

²⁶The contents of X can vary individual-by-individual, but its size does not.

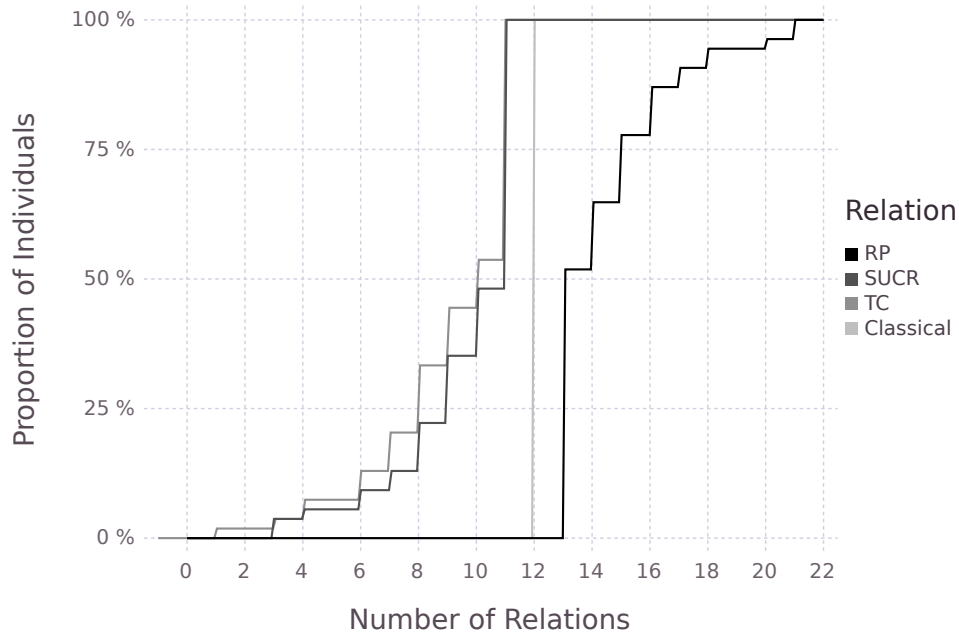


Figure 2.2: CDF of the number of relation elements at the individual level (for individuals with cyclic RP), in the experimental data.

SUCR and TC. To provide a stricter test of the completeness of these relations, we consider only relations over chosen bundles for our consumption data.

In this data, the average number of relation elements for SUCR is 7,076, which is 99.1% of the comparisons in a classical preference. The corresponding figures for TC are 7,021 and 98.3%.²⁷ The maximal number of relation elements is 7,135 for both SUCR and TC, which is very close to the size of a classical preference.

There is a sense in which the number of SUCR relation elements are over-counted in data sets without full observability, particularly when the relation contains cycles. It may be a reason to focus more on the completeness of SUCR in the experimental data, which has full observability or to focus instead on the completeness of TC in the consumption data. Additionally, this is a reason to look at the completeness of SUCR for those individuals with acyclic SUCR. Conditional on SUCR being acyclic, the average number of relation elements for SUCR is 7,084, which is 99.2% of the comparisons in a classical preference, and for the same subjects, the corresponding figures for TC are 7,043 and 98.6%.²⁸

²⁷The difference between the average number of relations for SUCR and TC is significant (the two-sided paired t-test p-value is <0.001). For a KS-test of equality of the distributions of SUCR and TC completeness, the p-value is <0.001 .

²⁸For subjects with acyclic SUCR, once again the difference between the average number of relations for SUCR and TC is significant (the two-sided paired t-test p-value is <0.001). For a KS test of equality of the distributions of SUCR and TC completeness, the p-value is <0.001 .

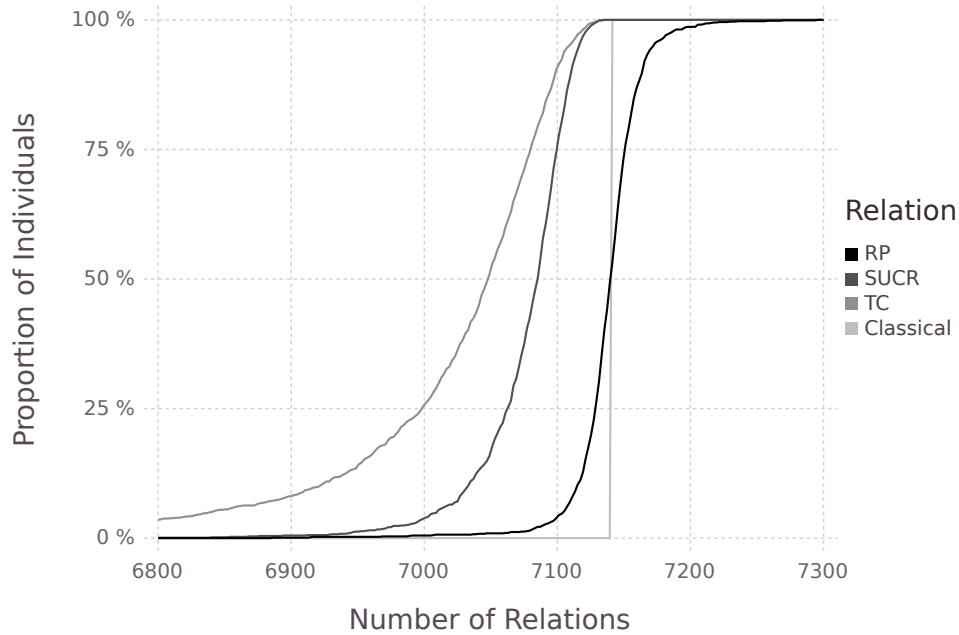


Figure 2.3: CDF of the number of relation elements at the individual level (for individuals with cyclic RP), in the consumption data.

2.3.4 Predictive Power of SUCR and TC

The completeness of SUCR and TC gives us a sense for the precision of their welfare guidance. However, it does not tell us exactly the precision of their welfare guidance for the observed choice sets, so we also calculate the predictive power of SUCR and TC “within sample” for both data sets.

To determine the predictions made by a relation for the observed choice sets, we follow Schwartz (1976) and Efe A. Ok (2002) in saying that the choice correspondence C induced by a (possibly incomplete) strict relation \succ is $C_\succ(Z) = \{x \in Z \mid y \succ x \text{ for no } y \in Z\}$.²⁹ The tightness of the predictions given by C is useful for studying the precision of welfare guidance because what is predicted to be selected from a choice set based on C_\succ is what is welfare optimal for that choice set. In the language of Bernheim and Rangel (2009), the elements of C_\succ are the “weak individual welfare optimum” of choice set Z .

A natural measure of the predictive power of C_\succ for an individual is the average size of C_\succ for that individual. For this measure, the highest possible predictive power corresponds to a value of 1 (a single alternative predicted from all choice sets), and larger values represent less predictive power (more alternatives predicted).

In the experimental data, SUCR predicts that an average of 1.32 alternatives could be chosen for

²⁹Bernheim and Rangel (2009) propose the same correspondence, which they denote as $m_\succ(Z)$.

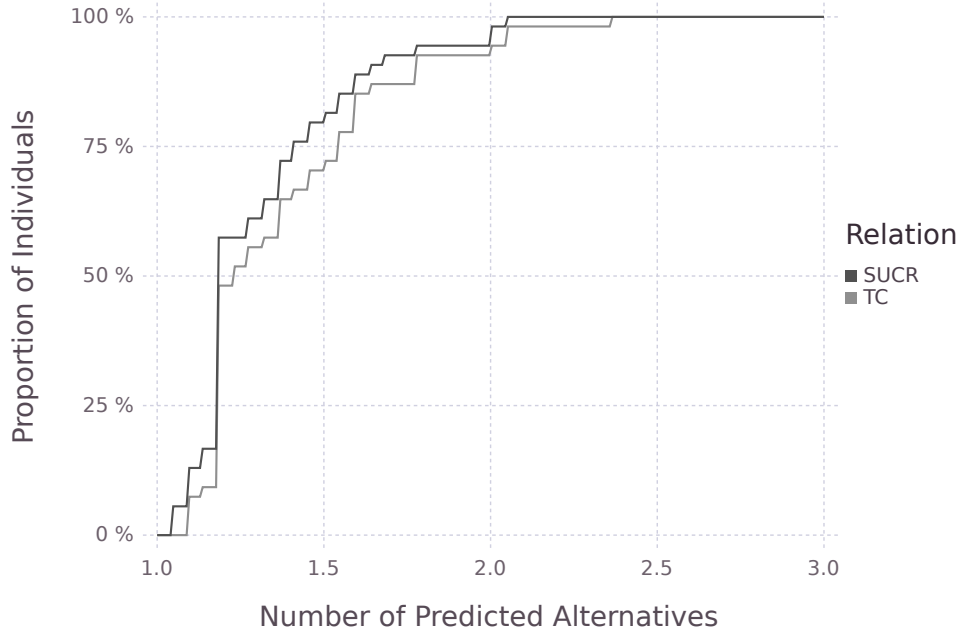


Figure 2.4: CDF of the average number of predicted alternatives for TC and SUCR at the individual level (for individuals with RP cycles), in the experimental data.

individuals with cyclic RP, whereas TC predicts that on average, 1.38 alternatives could be chosen.³⁰ In the consumption data, the average number of predicted alternatives is 1.33 for SUCR, and 1.65 for TC.³¹ Figures 2.4 and 2.5 shows that SUCR makes tighter predictions than TC at the individual level. For a KS-test of equality of the distributions of SUCR and TC for the average number of predicted alternatives at the individual level is 0.0095 on the experimental data and <0.001 on the consumption data.

As our primary measure of predictive power, we use the average value of Selten’s index (Selten (1991)) instead because it has a theoretical grounding, has been used for related questions in the literature, and accounts for the number of available alternatives. With Selten’s index, the proportion of choices that a theory predicts successfully within-sample is reduced by the “area”, which is the relative size of the predicted subset compared with the set of all possible outcomes. In the notation of Selten (1991)}, it is written as $m = r - a$, where r is the relative frequency of correct predictions and a is the area. Because SUCR and TC always successfully predict the chosen option, r is always equal to 1, but a can vary by relation and choice set. For a relation \succ and choice set Z , we define a as the proportion of alternatives that are predicted to be chosen, so that $a = \frac{|C_{\succ}(Z)|}{|Z|}$.

The choice set and its size are very straightforward to determine in the experimental data. As mentioned previously, in the consumption data we take the grand set of alternatives X to be the

³⁰This difference is significant, as the two-sided paired t-test p-value is 0.0025.

³¹This difference is significant, as the two-sided paired t-test p-value is <0.001 .

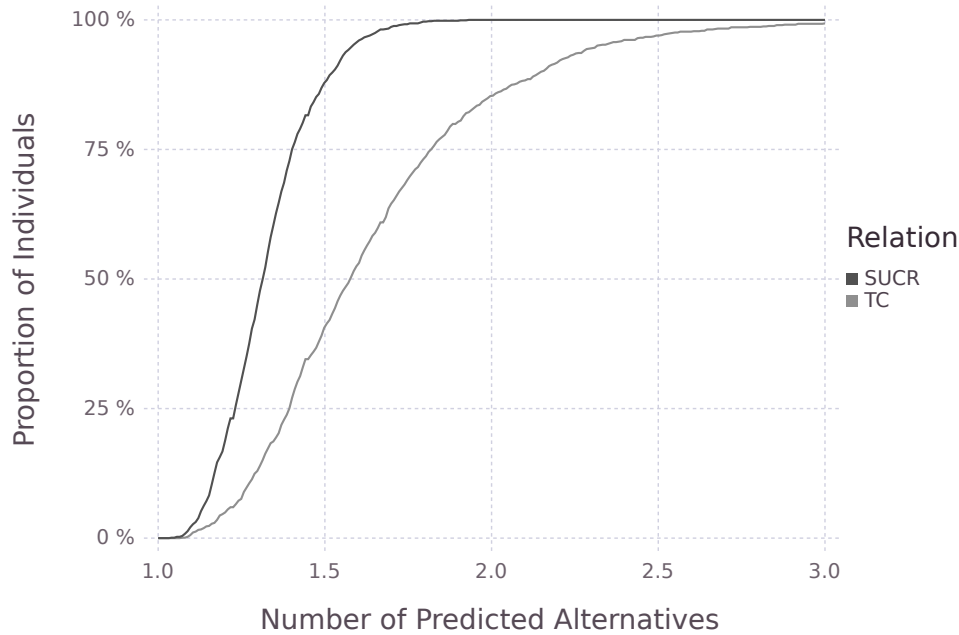


Figure 2.5: CDF of the average number of predicted alternatives for TC and SUCR at the individual level (for individuals with RP cycles), in the consumption data.

set of all bundles that an individual has chosen at some point (in order to provide a stricter test of completeness). Analogously, we take the choice set to be the set of all bundles that an individual has chosen at some point (all bundles in X) that are affordable at a given price and expenditure level.

We could consider the choice set to be every possible bundle on the budget line (as in Beatty and Crawford (2011)), but we consider this restricted space for three main reasons. First, it is far more computationally feasible to determine the set of predicted options given the large number of choices in our data set. Second, only bundles chosen elsewhere can generate inconsistencies. Third, this allows us to use the same metric across data sets. Fortunately, this approach provides a wide variety of choice set sizes, as shown in Appendix A.1.

One reason that we use Selten’s index is that it has an axiomatic foundation, and another is that it has been used elsewhere in the literature on empirical revealed preference analysis (Manzini and Mariotti (2010), Beatty and Crawford (2011), Dean and Martin (2016)). For example, Beatty and Crawford (2011) determine the fraction of demands that would pass a revealed preference test and then subtract this from an indicator for whether or not the observed choices passed the test. Their goal is to determine whether or not it is difficult for a set of choices to pass the revealed preference test for a given data set. Alternatively, Dean and Martin (2016) determine the average “distance” from rationality for all possible demands and then subtract this from the “distance” for observed choices.

Table 2.7: Average value of Selten’s index by choice set size in the experimental data (for individuals with RP cycles).

Choice set size	2	3	4	Average ^a
SUCR	0.40	0.53	0.60	0.46
TC	0.38	0.50	0.57	0.44
Classical	0.50	0.66	0.75	0.58

^a Average obtained by weighting by the number of sets.

In the experimental data, the average value of Selten’s index for SUCR is 0.46 for individuals with cyclic RP, and for TC it is 0.44. The difference between the average Selten’s index for TC and SUCR is significant, as the two-sided paired t-test p-value is equal to 0.0022. For the experimental data, the average theoretical maximum of Selten’s index is just 0.58, as shown in Table 2.7. For a set size of 2, the maximum value of Selten’s index is 0.5; for a set size of 3, it is 0.66; and for a set size of 4, it is 0.75.

In the consumption data, the average Selten’s index is 0.95 for SUCR and 0.94 for TC. The difference here is significant, as the two-sided paired t-test p-value is <0.001 . For the consumption data, the average theoretical maximum is 0.96.³² Selten’s index is higher in the consumption data compared to the experimental data because the size of the choice set appears in the denominator, and the choice set sizes are on average much higher on consumption data than they are on the experimental data.

Figures 2.6 and 2.7 provide the CDF of Selten’s index for both SUCR and TC for both data sets. The KS-test for equality of distributions has a p-value of <0.001 between SUCR and TC for both data sets.

As these results show, SUCR and TC have high predictive power on average for both the experimental and consumption data, which means that they provide precise welfare guidance. It qualifies our results on completeness shown in Section 2.3.3: the relations induced by SUCR and TC are complete enough to provide precise welfare guidance for observed choice sets.

2.3.5 Predictive Power and Revealed Preference Properties

In this section, we indicate when and why SUCR and TC have high predictive power and provide empirical evidence of these relationships. Specifically, we show that there are two properties of revealed preferences that are especially important for the predictive power of SUCR and TC: the of direct RP cycles and the *fraction* of RP cycles that are direct.

³²In order to compute the theoretical maximum of Selten’s index, we assume that in all choice sets exactly one alternative is predicted to be chosen, and then take the average over all choice sets.

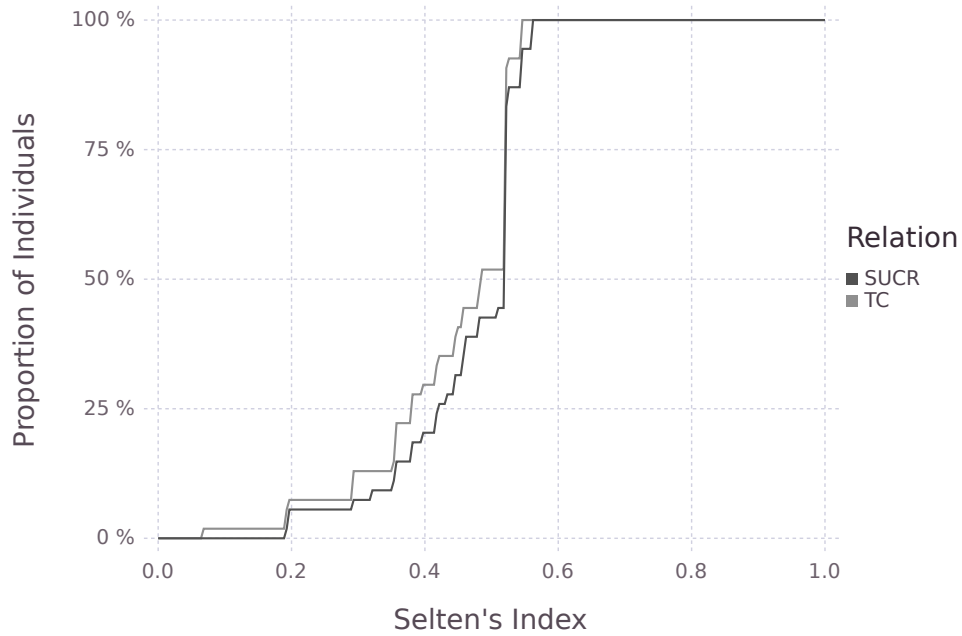


Figure 2.6: CDF of the average Selten's index at the individual level (for individuals with RP cycles), in the experimental data.

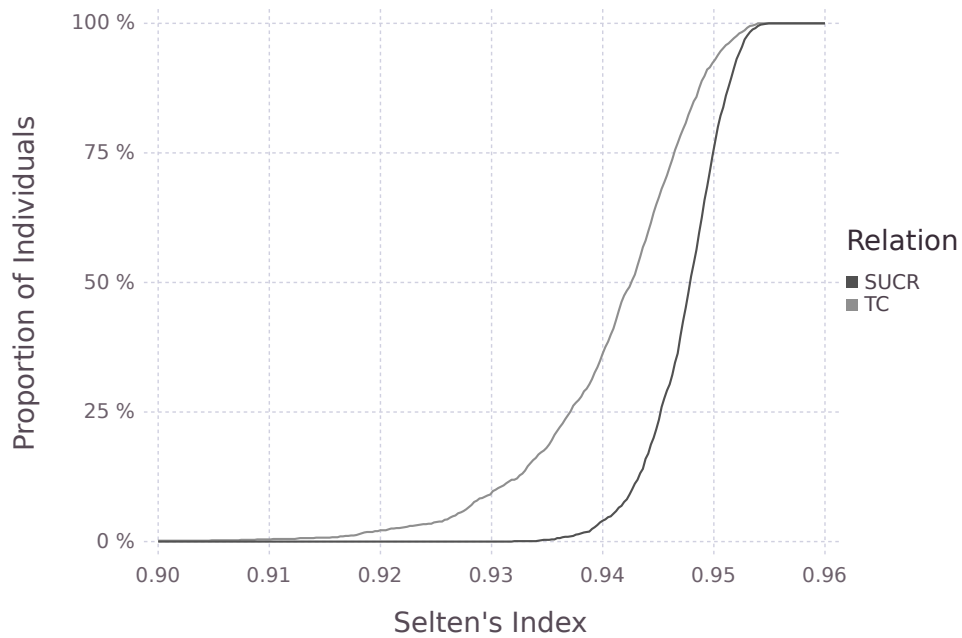


Figure 2.7: CDF of the average Selten's index at the individual level (for individuals with RP cycles), in the consumption data.

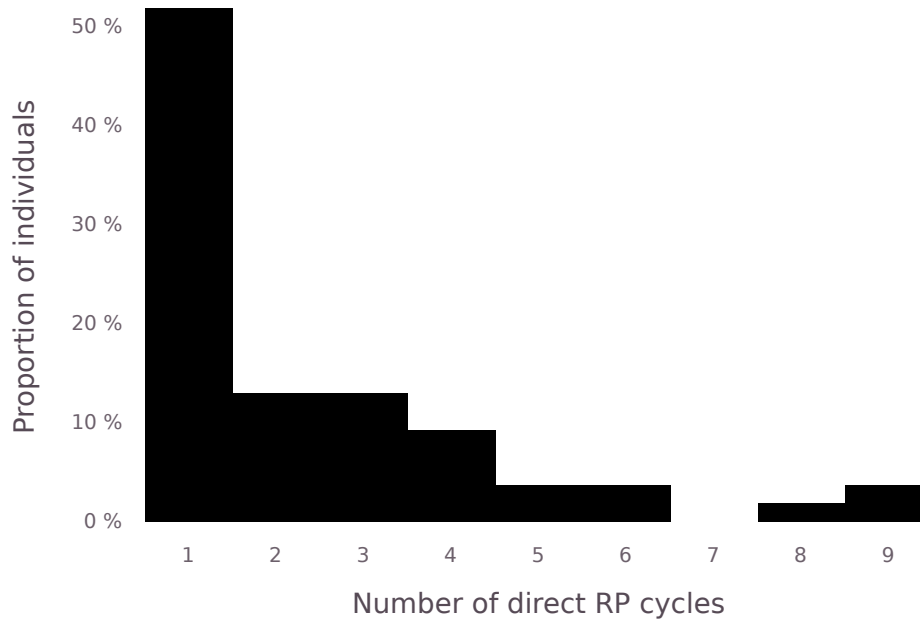


Figure 2.8: Histogram of the number of direct RP cycles per individual (for individuals with RP cycles), in the experimental data.

2.3.5.1 Number of Direct RP Cycles

As discussed previously, SUCR and TC remove all revealed preference relation elements that produce direct RP cycles, so their ability to make precise predictions should be linked to the number of such cycles. For individuals with RP cycles, the average number of direct RP cycles is 2.43 in the experimental data, and 31.28 in the consumption data.³³ Figures 2.8 and 2.9 show the distribution of the number of direct RP cycles for individuals with RP cycles, which indicates there is heterogeneity in the number of direct RP cycles. A majority of individuals have very few such cycles, but some exhibit comparatively more cycles.

Table 2.8 provides a summary of the correlations with the average value of Selten's index at the individual level. As expected, the predictive power is highly and negatively correlated with the number of direct RP cycles. The more direct RP cycles there are, the fewer relations there will be in TC and SUCR, and therefore the less predictive power they will have. Looking at correlation with Selten's index, we find that the number of direct RP cycles is more highly correlated than the number of RP cycles of length three or even the total number of cycles of length 2, 3, and 4.

³³To avoid double-counting, xPy and yPx count as a single cycle. Also, when multiple choices reveal that x is preferred to y , then we count the relation element just once when counting cycles. For instance, if two choices reveal xPy and one choice reveals yPx , then this counts as just one direct RP cycle

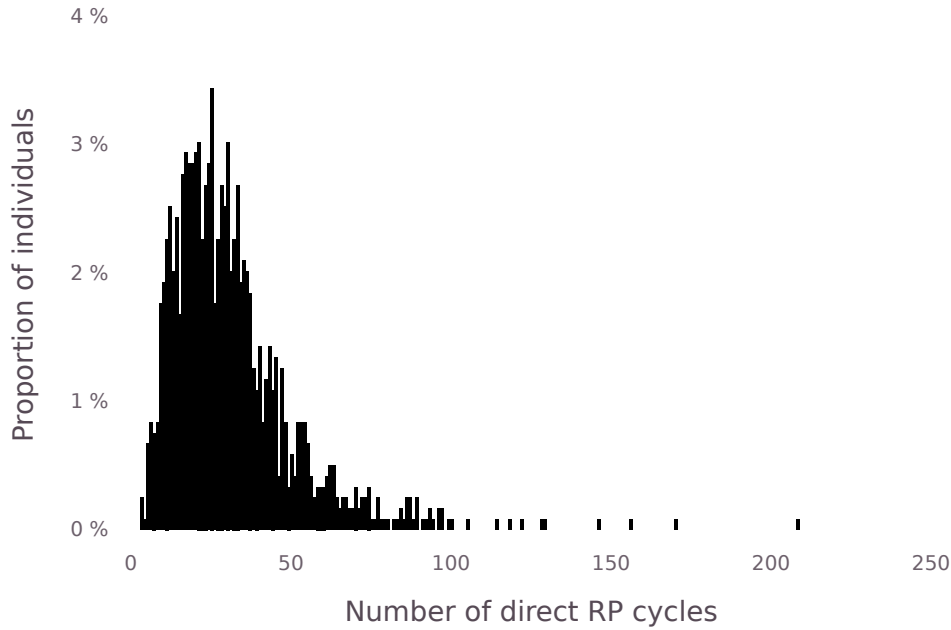


Figure 2.9: Histogram of the number of direct RP cycles per individual (for individuals with RP cycles), in the consumption data.

2.3.5.2 Fraction of RP Cycles that are Direct

In addition to the number of direct RP cycles, if there are many cycles of longer length relative to the number of direct RP cycles, then this is a problem for both SUCR and TC. For one, SUCR will be cyclical if RP cycles remain after ignoring RP relation elements that generate direct RP cycles, so SUCR is more likely to be cyclical if there are many cycles of longer lengths relative to the number of direct RP cycles. Besides, TC ignores RP relation elements beyond just the RP relation elements that produce direct RP cycles, so its predictive power is likely to be lower if there are many cycles of longer lengths relative to the number of direct RP cycles.

Also, the fraction of RP cycles that are direct has a possible interpretation in terms of behavioral welfare. It could be argued that direct inconsistencies should be more obvious to decision makers, so are more likely to represent intentional violations. As such, this fraction could be interpreted as the fraction of RP cycles that represent “real” behavioral preference changes.

However, determining the fraction of RP cycles that are direct requires determining the number of cycles of all lengths, which can become computationally burdensome. Instead, we analyze an upper bound on this fraction that is much quicker and easier to calculate: the ratio of the number of length two cycles to the number of length two and length three cycles. We call this measure the “directness index” (DI henceforth).

Figures 2.10 and 2.11 show the histogram of DI for individuals with RP cycles. For the experimental

Table 2.8: Summary of correlations with the average Selten's index at the individual level (for individuals with RP cycles).

	Selten's Index					
	Experimental		Consumption		Acyclic SUCR	
	SUCR	TC	SUCR	TC	SUCR	TC
Number of relation elements	0.98	0.92	0.59	0.87	0.58	0.82
Number of direct RP cycles	-0.98	-0.92	-0.75	-0.81	-0.75	-0.78
Number of length 3 RP cycles	-0.89	-0.92	-0.49	-0.66	-0.48	-0.60
Number of length 2, 3, and 4 RP cycles	-0.92	-0.92	-0.35	-0.53	-0.31	-0.43
Directness index (DI)	0.58	0.78	0.56	0.69	0.52	0.65

data, there is a significant mass point at 1. The distribution for the consumption data (Figure 2.11) is hump-shaped around .5, but there is also a mass point at 1. In the experimental data, the average DI is 0.79 for individuals with RP cycles, and the average is 0.58 if we exclude individuals with a DI of 1. In the consumption data, the average DI is 0.50, and it falls to 0.49 if we exclude individuals with a DI of 1.

The correlations with the predictive power are given in Table 2.8. As expected, DI is highly and positively correlated with the value of Selten's index for TC. In the experimental data, the correlation is 0.78, and in the consumption data, the correlation is 0.69.

We find that as expected, the acyclicity of SUCR in the consumption data is highly and positively correlated with DI. The correlation between DI and a dummy variable that takes a value of 1 when SUCR has no cycles is -0.38. This correlation is even higher than the correlation between this dummy variable and the number of direct RP cycles, which is 0.36.

2.3.5.3 Regression Analysis

Because DI is highly correlated with the number of direct RP cycles (-0.68 in the experimental data, and -0.76 in the consumption data), we also use regressions to examine the impact of DI while controlling for the number of direct RP cycles. As shown in specification 2 of Table 2.9, DI is positively and significantly related to Selten's index, even when controlling for the number of direct RP cycles.

2.4 Discussion and Conclusion

In this paper, we provide the first non-parametric empirical application of SUCR and the first empirical application of TC. The resulting analysis helps to provide an empirical answer to whether a

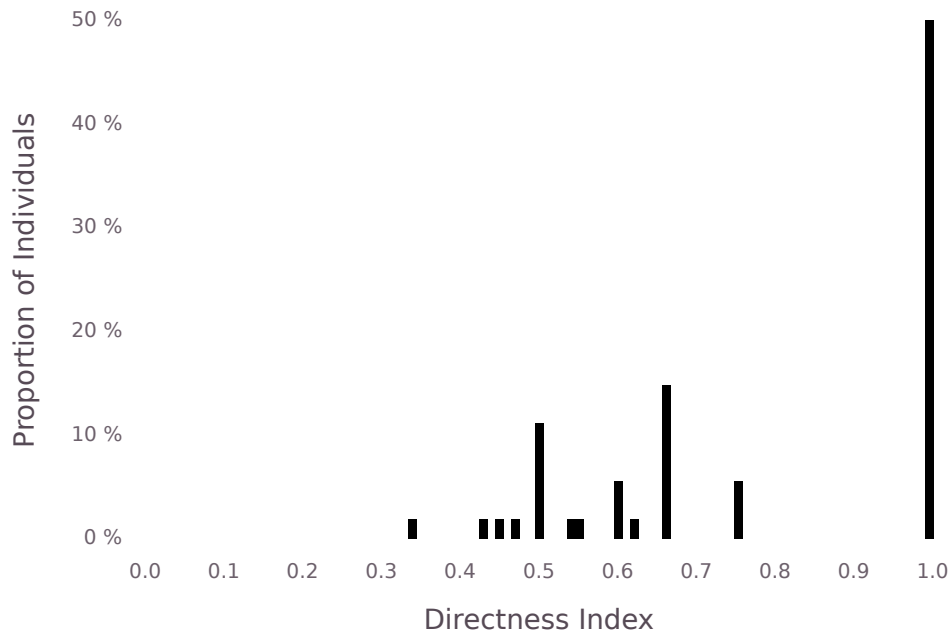


Figure 2.10: Histogram of the directness index (DI) at the individual level (for individuals with RP cycles), in the experimental data.

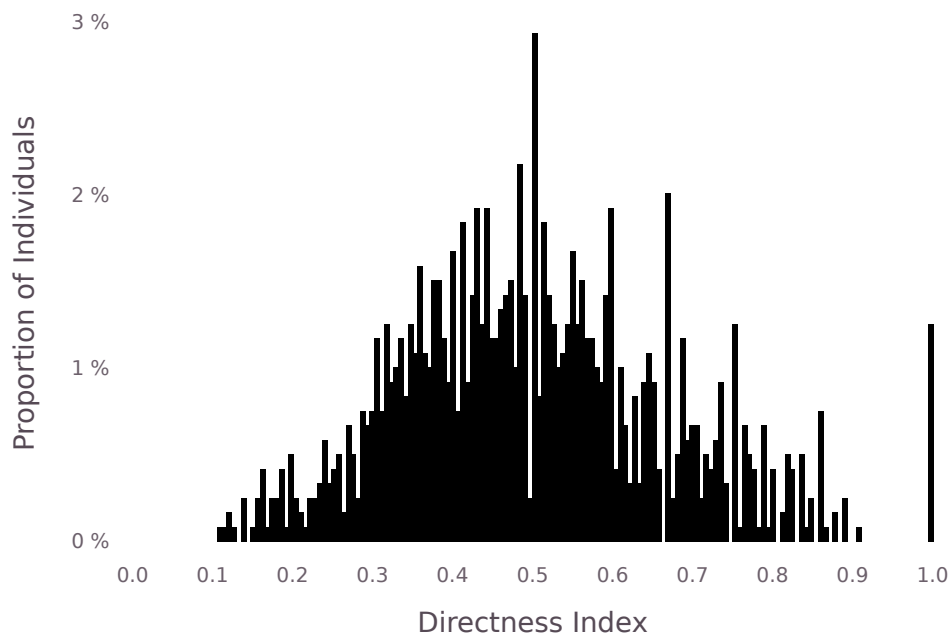


Figure 2.11: Histogram of the directness index (DI) at the individual level (for individuals with RP cycles), in the consumption data.

Table 2.9: Regressions of the average Selten’s index for TC onto the number of direct RP cycles and the directness index (DI).

	Selten’s Index		
	Experimental	Consumption	Acyclic SUCR
Number of direct RP cycles	-0.038*** (0.005)	-0.00026*** (0.00003)	-0.00027*** (0.00002)
Directness index (DI)	0.144** (0.046)	0.00892** (0.00201)	0.00513*** (0.00141)
Observations	54	1,190	946
Adjusted R ²	0.8872	0.6723	0.6159

Note:

Robust standard errors in parenthesis. Significance levels: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

substantial model structure is needed to give precise welfare guidance when choices appear inconsistent due to behavioral biases.

For both data sets considered in this paper, we find that SUCR is most often acyclic and that both SUCR and TC have high predictive power, which means that they offer precise welfare guidance. Of course, to provide a more comprehensive and general answer to when SUCR is acyclic and has high predictive power, we would need to look at other experimental and non-experimental data sets, such as those examined in the behavioral economics literature.

It should be noted that neither of our data sets were cherry-picked to produce a desired result. Our prior belief was that SUCR and TC would not offer precise welfare guidance for the data sets we examine in this paper. Instead, our results lead us to conclude something entirely different for these data sets.

We feel that the data sets examined in this paper represent valid test sets because behavioral biases are likely to influence choices made in these settings. For instance, the alternatives presented at the top of the list in the experiment are likely to be picked more often. Also, in grocery store purchases, consumers may be drawn to a product due to its position in the aisle or special display case. Alternatively, individuals may be tempted to buy products that they do not want because they are hungry.

Finally, while the original formulation of SUCR was given for choices from menus, we felt it was important to examine the performance of SUCR for choices from budget sets. Not only are choices from budget sets a canonical revealed preference data set, but in many applications, there are prices associated with goods, so to ignore budget set data is to ignore many real-world settings.

Chapter 3

Identifying Choice Correspondences

3.1 Introduction

In a seminal paper, Samuelson (1938) introduced the revealed preferences method. He linked preferences and choices by positing that chosen alternatives are better than unchosen ones, thus revealing the preference of the decision maker. In experiments and most real-life settings, we identify a *choice function*, using a forced single choice method: the decision maker chooses exactly one alternative from the choice set. In principle, however, choices are commonly modeled with a *choice correspondence*: the decision maker chooses a non-empty set from the choice set.¹

In this chapter, we introduce *pay-for-certainty*, a method for identifying choice correspondences in experiments in an incentive-compatible manner.² We allow decision makers to choose several alternatives, provide a small incentive for each alternative chosen, and then randomly select one for payment. Some experiments have allowed decision makers to choose several alternatives, but none have tried to elicit the choice correspondence of decision makers in an incentive-compatible manner. Danan and Ziegelmeyer (2006) have allowed decision makers to postpone their choice, at a cost. Agranov and Ortoleva (2017) made subjects knowingly repeat their choice and allowed them to switch. Ong and Qiu (2018), Sautua (2017), and “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019) have allowed them to delegate their choice to a random device, with and without costs. Finally, Costa-Gomes, Cueva, and Gerasimou (2019) have allowed decision makers to postpone their choice, at a cost. All these methods do not make choosing all the maximal alternatives the dominant strategy. At best, it is a weakly dominant strategy. They may fail to identify the choice correspondence of decision makers if they

¹One real-life example of the choice of a non-empty set from the choice set is approval voting. Decision makers can vote for all the candidate they deem acceptable.

²Azrieli, Chambers, and Healy (2018) provide a guide in what is incentive-compatible in experiments. It relies in particular on only one choice being paid, but we will not dwell on the practice of incentive-compatibility here, this discussion is postponed to Chapter 4.

are indifferent between two or more alternatives.

Identifying choice correspondences is relevant to welfare analysis. Danan, Gajdos, and Tallon (2013) and Danan, Gajdos, and Tallon (2015) have shown that the aggregation of incomplete preferences might be relatively more straightforward than the aggregation of complete preferences. Maniquet and Mongin (2015) have shown that it is possible to aggregate preferences with no more than two indifference classes. In order to study indifference or incomplete preferences, we need a choice correspondence, as shown in Mandler (2005), Eliaz and Efe A. Ok (2006) and Aleskerov, Bouyssou, and Monjardet (2007). The following example illustrates why we want to identify a choice correspondence, rather than a choice function.

Example 3.1 (Pizzeria). On a small island, there is a pizzaiolo who produces three kinds of pizzas, a vegetarian pizza (V), a four-cheese pizza (C) and a ham pizza (H). He is the sole pizzaiolo on the island, so there are no competitive forces on the pizza market. The preferences of the islanders, which he does not know, are as follows:

- Half of the islanders are indifferent between the vegetarian and the four-cheese pizza, and prefer both to the ham one, i.e., $V \sim C \succ H$;
- Another half is indifferent between the ham and the four-cheese pizza, and prefer both to the vegetarian one, i.e., $H \sim C \succ V$.

For cost reasons, the pizzaiolo wants to produce only one kind of pizza. The natural choice to make in his situation is to keep the most chosen pizza. Is it the best choice for the welfare of all islanders?

Let us say that from the first half of the population (with preference $V \sim C \succ H$), a proportion p chooses the vegetarian pizza. From the second half of the population (with preference $H \sim C \succ V$), a proportion q chooses the ham pizza. The summary of the probability of each pizza being chosen is given in Table 3.1

Table 3.1: Fraction of the population choosing each pizza, with $0 \leq p, q \leq 1/2$.

	Vegetarian	Ham	Four-Cheese
Observed Choice	p	q	$1 - p - q$
Maximal Alternatives	$\frac{1}{2}$	$\frac{1}{2}$	1

Figure 3.1 shows the values of the (p, q) for which each kind of pizza would be chosen. Overall, in two-thirds of the situations, the pizzaiolo would keep the four-cheese pizza. In one-third of the situations, he will keep another pizza, which would decrease the welfare of the population. If islanders use a coin toss to decide when they are indifferent, and the number of islanders coming to the pizzeria is large enough, he will keep the four-cheese pizza. If the number of observations is

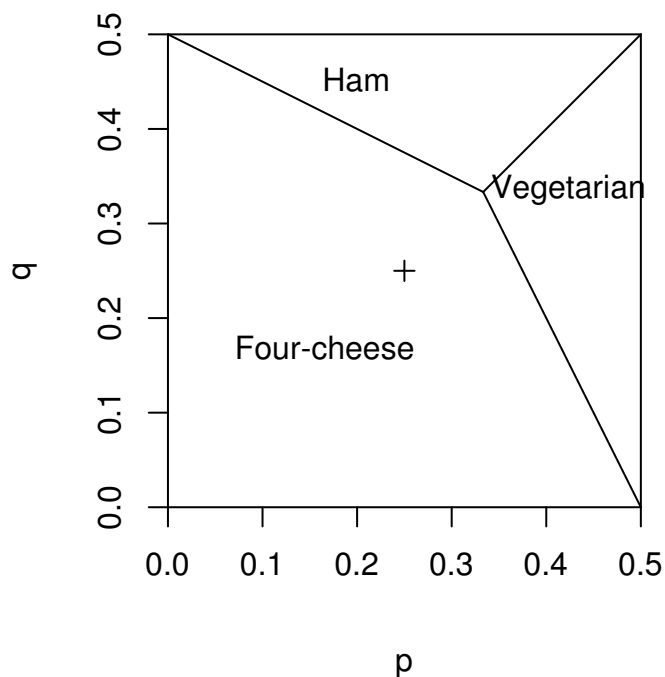


Figure 3.1: In which probability pair is each pizza the most chosen?

quite small, on the other hand, he may keep the vegetarian or the ham pizzas. Had the pizzaiolo known the preferences of the islander, he would have kept the four-cheese pizza in all cases. It is the pizza that does not decrease welfare for any islander, so it is the best to keep from a collective welfare standpoint.

This example shows that in general, a proper revelation of preferences and the subsequent study of welfare require more than identifying a choice function. It requires the set of maximal alternatives, i.e., the choice correspondence for all decision makers. Besides, identifying a choice correspondence for each decision maker allows us to explore new questions. First, as explained in Section 1.2, it allows us to go beyond the classical revealed preference paradigm. Second, it allows us to study the full extent of indifference in choices.

The main challenge to overcome is that choice correspondences may not be obtained directly. A simple direct method would allow decision makers to choose several alternatives and then randomly select one for payment. This method might not identify their choice correspondence in the presence of indifference. Indeed, choosing all maximal alternatives is not a strictly dominant strategy in such settings.

In Example 3.1, let us imagine now that the pizzaiolo allows all clients to choose as many pizzas

as they wish. If they choose more than one pizza, he tells them they will get one of their chosen alternatives, chosen by him at random presumably. What will clients with preferences $V \sim C \succ H$ choose? They can choose either $\{V, C\}$, $\{C\}$ or $\{V\}$ and get the same satisfaction. Because choosing all the maximal alternatives is not a strictly dominant strategy, it creates a uniqueness problem for the observer – the pizzaiolo. This procedure does not guarantee that the decision maker chooses all maximal – i.e., Pareto-superior – alternatives. On the other hand, revealed preference models based on choice correspondences (see, for instance, Amartya Sen (1971), Schwartz (1976), Nehring (1997)) assume that the chosen set is the set of maximal alternatives. Models studying relaxations of the classical paradigm, in particular, relies on this interpretation of choice correspondences.

Pay-for-certainty solves the uniqueness problem by incentivizing decision makers to choose *all* maximal alternatives. For each alternative chosen, the decision maker earns an *additional payment* $\varepsilon > 0$. In the previous example, the clients are better off by choosing $\{V, C\}$ and getting the vegetarian or the four-cheese pizza *and* 2ε in additional payment rather than choosing and getting $\{V\}$ or $\{C\}$ and ε . The full characterization of *pay-for-certainty* is thus:

In each set of alternatives S , the decision maker chooses all the alternatives he wishes, i.e., a subset $c(S)$. He earns a (small) bonus payment of $\frac{|c(S)|}{|S|}\varepsilon$ by chosen alternatives. The alternative he gets is picked from her chosen alternatives using a selection mechanism, for instance, a uniform random draw.

The bonus payment of pay-for-certainty implies forgone gains when the chosen set is not the whole set. Choosing only one alternative earns $\frac{1}{|S|}\varepsilon$, whereas choosing two earns $\frac{2}{|S|}\varepsilon$, and so on. We show that under mild monotonicity conditions, the decision maker chooses all maximal alternatives. The choices observed with pay-for-certainty and a ε bonus payment is called the ε -*correspondence* hereafter.

A possible downside is that some chosen alternatives might not be maximal if ε is large and the differences among the (direct) payoffs of some alternatives are within $\frac{1}{|S|}\varepsilon$ of each other. The method shares many features with the experiment of Costa-Gomes, Cueva, and Gerasimou (2016), but there are some key differences.³ First, choosing several alternatives implies a gain, not a loss. Second, the gain here will be much lower in magnitude, theoretically eliciting more indifference relations. Last, the dominant strategy of a decision maker who is indifferent between different alternatives is to choose all of them.

We provide two identification results in this chapter. The first one is a limit identification result, showing that if ε is small enough and the grand set of alternatives is finite, we identify the choice correspondence. The second result is a set inclusion identification result. It says that in any

³The experiment uses the theory built in Gerasimou (2017) to identify indecision, by using choice correspondence that might be empty valued. The purpose is a little different from our purpose here, which help explains why we reach different mechanisms.

choice set, the 0-correspondence is included in the choice correspondence, itself included in the ε -correspondence. When this is satisfied, the choice correspondence of the decision maker is *partially* identified. Precise identification of the maximal alternatives in a given set happens when the choices with no bonus payment and a positive bonus payment are equal. When precise identification happens for all choices, the choice correspondence is *fully* identified. The fully identified choice correspondence of a decision maker is his theoretical choice correspondence.

In practice, the second result is directly testable, whereas the first one is not. In general, it is impossible to guarantee full identification for everyone. We, therefore, provide results for the compatibility of different decision-making models when the choice correspondence is partially identified.

This chapter is structured as follows. Section 3.2 describes the different feature of the pay-for-certainty method. Sections 3.3 focuses on pay-for-certainty with a uniform selection mechanism and provides its domain of validity and identification results. Section 3.3.3 studies some models that can be falsified under partial identification.

3.2 Definition of the Pay-For-Certainty Method

This section introduces, describes, and discusses the different feature of the pay-for-certainty method.

3.2.1 General Setup

Formally, we use the same setup as in the introduction, but we limit ourselves to finite data. Now, X is a *finite* set of alternatives.⁴ S designs a non-empty subset of X , a *choice set*. Choices of decision makers are modeled with a *choice correspondence*, as in Definition 1.1.

Alternatives in $c(S)$ are the ones *chosen* in S , whereas alternatives in $S \setminus c(S)$ are *unchosen* in S . For incentive purposes, the difference between chosen and unchosen alternatives is that the decision maker never gets unchosen alternatives, whereas he gets *one* of the chosen alternatives. In this sense, the decision maker wants alternatives from $c(S)$ and does not want alternatives from $S \setminus c(S)$. We are looking to use strict revealed preferences, rather than weak revealed preferences, as the former has more explanatory power and is more stringent than the latter. Appendix C.2 quantifies this difference in a particular case.

Assumption 3.1 (Preferences). *The decision maker has preferences \succeq that are at least a partial order. That is, \succeq is quasi-transitive.*

⁴The design does not extend readily to infinite sets. We discuss why in Appendix B.2.

Importantly, preferences do not have to be transitive, but only quasi-transitive. In particular, the indifference part does not have to be transitive, which is essential to explore intransitive indifference. Saying that chosen alternatives are maximal does not adequately characterize chosen and unchosen sets. Indeed, at least two competing assumptions on the choice correspondence are possible. Call $M(S)$ the set of maximal alternatives.

Definition 3.1 (Set of Maximal Alternatives in S ($M(S)$)).

$$M(S) = \{x \in S \mid \text{there is no } y \in S, y \succ x\}$$

The set of *maximal* alternatives in S is the set of all alternatives which are not strictly worse than any other alternative.

As long as S is non-empty, $M(S)$ is non-empty, as the set S is finite and \succ is transitive. $M(S)$ contains all the alternatives a maximizing decision maker potentially chooses. All else equal, it is sub-optimal for the decision maker to choose an alternative in $S \setminus M(S)$, as he could choose a better alternative. That is, the set of chosen alternatives is a subset of the set of maximal alternatives, i.e., $c(S) \subseteq M(S)$. In that case, we can only assume that the revealed preferences are *weak*.

Most theoretical work (see Schwartz (1976), Aleskerov, Bouyssou, and Monjardet (2007), Amartya Sen (1997), Nehring (1997), among others) adopt the following stronger statement. The set of chosen alternatives *is* the set of maximal alternatives, i.e., $c(S) = M(S)$. This stronger assumption implies that any unchosen alternative in S is dominated by another alternative in S , and by transitivity and finiteness of S , by an alternative in $M(S)$:

$$\text{for all } x \in S \setminus M(S), \text{ there exists } y \in M(S), y \succ x$$

It implies that choices in binary sets reveal the *strict* preferences of the decision maker, i.e., $c(S) = M(S)$. If we observe that $c(\{x, y\}) = \{x\}$, we can say for sure that $x \succ y$.

While very convenient, the strict revelation assumption is by no means guaranteed to hold in practice. The remainder of this section studies the conditions under which this assumption on chosen alternatives is legitimate with pay-for-certainty. That is, under which assumptions can we say that $c(S) = M(S)$. To be consistent with the introduction, we will note from now on $c(S)$ for $M(S)$.

3.2.2 Definition of the *Pay-For-Certainty* Method

The objective of pay-for-certainty is to recover the set of maximal alternatives of a decision maker in a set S . In most experiments, decision makers are forced to choose a single choice. For each choice

set, decision makers choose precisely one alternative, which is then given to them. It is perfect for identifying the choice function of the decision maker. Arguably, this is close to the situation in the field, where decision makers generally choose one and only one alternative. There is one key difference, however. In the field, it is generally possible to postpone the choice, which is rarely the case in experiments. Dhar and Simonson (2003) have shown that forcing choice modifies the choice of decision makers. Danan and Ziegelmeyer (2006) and Costa-Gomes, Cueva, and Gerasimou (2019) show experimentally that decision makers value the possibility to postpone their choice. Agranov and Ortoleva (2017) show that sometimes, decision makers are not sure of which alternative is the best in a choice set. These pieces of evidence imply that decision maker like some flexibility in their choices. Choice correspondences introduce this flexibility in experiments by not forcing decision makers to select precisely one alternative.

There are two keys to incentivize decision makers to choose exactly their set of maximal alternatives. First, we must incentivize them to choose all their maximal alternatives. Second, they must not choose *more* than their maximal alternatives. The *selection mechanism* takes this second role.

3.2.2.1 Incentive to Choose Several Alternatives

Contrary to choice functions elicitation, there are different ways to elicit a choice correspondence. The simplest was introduced in Example 3.1. Call it the *0-correspondence elicitation*:

In every choice set S , the decision maker chooses a non-empty subset $c(S)$. A selection mechanism selects the alternative.

The 0-correspondence elicitation procedure does not guarantee that decision makers choose the set of best alternatives. This non-maximal problem is very general and arises for a classical decision maker because of indifference. It can only guarantee that $c(S) \subseteq M(S)$. One solution to elicit indifference dates back to Savage (1954) and was formalized by Danan (2008): costly strict preferences. If a decision maker is indifferent between two alternatives x and y , then any small gain (cost) added to one alternative will tip the choice in its direction (the opposite direction).

Adding a small gain for each alternative chosen incentivizes the decision maker to choose larger sets when he is indifferent between alternatives. We build the *pay-for-certainty* method on this intuition.

From the choice set S , each alternative chosen adds a *bonus payment* of $\frac{1}{|S|}\varepsilon > 0$ per alternative to the gain of the decision maker.⁵ The total additional payments are $\frac{|c(S)|}{|S|} \times \varepsilon$. The alternative he gets is selected using a selection mechanism.

⁵The gain or loss does not have to be monetary. It only has to be perceived as a cost or a gain to be used as payment. Time, for instance, could be used. The payment would then increase or decrease the time spent in the laboratory.

The selection mechanism is not *a priori* specified, but in practice, in Chapter 4, we use the uniform selection mechanism. The introduction of the payment breaks indifference but comes at a cost. If a decision maker slightly prefers x to y , and his preference is so weak that the difference is hardly perceptible, he will choose $\{x, y\}$ and we will think he is indifferent between x and y , which is not the case. If ε is large enough, choosing $\{x, y\}$ and getting ε is better than choosing $\{x\}$ (or $\{y\}$) and getting $\frac{\varepsilon}{2}$. Pay-for-certainty might bundle some strict preferences with indifference. In theory, this problem vanishes when ε tends to zero. In practice, ε cannot be vanishingly small, and the problem might persist. The error made is, by construction, bounded above in monetary terms by ε .

The bonus payment for each alternative chosen depends on the size of the choice set. Another possible incentive mechanism is to use a linear bonus payment, which is more straightforward than a proportional one to explain. We find, however, that the proportional bonus payment has better incentive properties in a simple case, as explained in Section 3.2.2.3.

One benefit of pay-for-certainty for indifference elicitation is that the payment is not directly related to the alternative. Indeed, one way to implement indifference is, following Danan (2008), to observe choices where each alternative has a small bonus payment associated with it, in turns. So in the set $\{x, y\}$, we would need two choices: $\{(x + \varepsilon), y\}$ and $\{x, (y + \varepsilon)\}$. One loss, however, is to lose the direct identification of intransitive indifference (or incompleteness): observing that in the first set, the whole set is chosen and in the second set, the whole set is also chosen cannot be explained with (transitive) indifference. It is, however, quickly hard to implement when X becomes large.

3.2.2.2 Selection Mechanisms

With choice correspondences, for incentive purposes, precisely one of the selected alternative must be used for the payment. We select this alternative through a *selection mechanism*. It associates to a set of alternatives one alternative from this set. The first selection mechanism we introduce is the *uniform selection mechanism*.

Selection Mechanism: *Uniform* When the chosen set contains more than one alternative, the decision maker gets the alternative drawn using a uniform random draw over the set of chosen alternatives.

The likelihood of getting a chosen alternative is $\frac{1}{|c(S)|}$ with the uniform selection mechanism. Adding an alternative in the chosen set has two consequences: it is now possible to get this alternative, *and* it decreases the chances of getting other chosen alternatives. We will study this selection mechanism in more details later. Other selection mechanisms are possible, and some of them have been used in the literature.

Selection Mechanism: *Own Randomization* When the chosen set contains more than one alternative, the decision maker chooses a distribution on chosen alternatives. The decision maker gets the alternative drawn from her distribution on chosen alternatives.

In this mechanism, the decision maker chooses the distribution. It is more complicated to implement experimentally, as it adds more features to the elicitation procedure, but yields to potentially more fruitful results. Some experimental results about own randomization are given in Appendix B.1.

Another possible mechanism is delegation – to someone else.

Selection Mechanism: *Delegation* When the chosen set contains more than one alternative, the decision maker gets an alternative selected by someone else.

In this mechanism, the selection process is now a black box for the decision maker, and potentially the observer. They do not know how the selection is made. Some assumptions are required to ensure correct elicitation. The first one is the absence of preference for delegation when preferences are strict. If subjects prefer to delegate no matter what happens, they will choose the whole set. The second one is the absence of an “experimenter” effect. The beliefs on the selection mechanism might influence the choices made by the decision maker. The uniform selection mechanism explicitly states the probabilities of getting an alternative. As long as the decision makers trust the experimenter, the beliefs should be correct and commonly known. With delegation, the constraints on the beliefs on the selection mechanism are hard to infer for the observer. For instance, decision makers might believe that the choosers will behave adversely, by choosing, for instance, the least costly alternative to the experimenter. One way to mitigate that might be to delegate anonymously to other subjects in the experiment and not to the experimenter. The point is, beliefs are hard to control with delegation, and the observed choices might reflect both the preference and the beliefs of the decision makers.

The last selection mechanism is flexibility.

Selection Mechanism: *Flexibility* When the chosen set contains more than one alternative, the decision maker will face the chosen set again later and will have to select one alternative.

This mechanism has been used in at least two experiments, one by Danan and Ziegelmeyer (2006) and the other by Costa-Gomes, Cueva, and Gerasimou (2019). It can be used with the pay for certainty choice procedure, under the condition that payments are costs and not gains. Otherwise, the incentive scheme is such that choosing the whole set and then selecting the alternative in the second choice is a strictly dominant strategy. With a cost, the incentive is to reduce as much as possible the chosen set in the first stage. It is not possible to guarantee that decision makers always choose the set of maximal alternatives with flexibility, and thus indifference may be underestimated.

3.2.2.3 Linear versus Proportional Bonus Payment

The justification behind the choice of a proportional rather than a linear payment scheme is quite simple. Imagine the choice between two alternatives x and y , with a linear payment scheme. That is, for each alternative chosen, the bonus payment is of ε , no matter the size of the choice set. Let us assume, for simplicity, that the decision maker maximizes a quasi-linear expected utility. The selection mechanism is a uniform distribution over the set of chosen alternatives. That is, he values the set $\{x, y\}$, $u(\{x, y\}) = \frac{x+y}{2} + 2\varepsilon$. By a slight abuse of notation, because of quasi-linear expected utility, we note x the utility of x .

Let us also assume that he prefers x to y , that is $x \succ y$. Clearly, the decision maker will never choose y , as $u(\{x\}) > u(\{y\})$. Now, the choice between $\{x\}$ and $\{x, y\}$ depends on ε :

$$\begin{aligned} u(\{x\}) &> u(\{x, y\}) \\ \Leftrightarrow x + \varepsilon &> \frac{x+y}{2} + 2\varepsilon \\ \Leftrightarrow x - y &> 2\varepsilon \end{aligned}$$

So $\{x\}$ is chosen when x 's utility is higher than y 's utility and 2ε .

Introduce now an additional alternative z . Introducing a third alternative z should not change the choice between x and y . That is, if $\{x\}$ is chosen over $\{x, y\}$, then $\{x, z\}$ should be chosen over $\{x, y, z\}$.

$$\begin{aligned} u(\{x, z\}) &> u(\{x, y, z\}) \\ \Leftrightarrow \frac{x+z}{2} + 2\varepsilon &> \frac{x+y+z}{3} + 3\varepsilon \\ \Leftrightarrow x + z - 2y &> 6\varepsilon \\ \Leftrightarrow \frac{x+z}{2} - y &> 3\varepsilon \end{aligned}$$

Now if z is, in fact, a copy of x , then we have in the second that x 's utility needs to be higher than y 's utility and 3ε , not 2ε . It is quite easy to see that a proportional payment does not have this feature. In general, a proportional payment has the benefit of taking into account the variation in probabilities induced by the uniform selection mechanism. A linear payment scheme cannot capture this externality.

In the next sections, we will study *under which condition the pay for certainty method with the uniform selection mechanism allows one to recover the set of best alternatives*. We need two kinds of assumptions: assumptions on preferences and assumptions on behavior.

3.2.3 Pay-For-Certainty with a Uniform Selection Mechanism

We focus on one pay-for-certainty, *pay-for-certainty with the uniform selection mechanism*.

From the choice set S , each alternative chosen adds a *bonus payment* of $\frac{1}{|S|}\varepsilon > 0$ per

alternative to the gain of the decision maker. The total potential bonus payments in a set is ε . The alternative he gets is selected using a uniform random draw on the set of chosen alternatives $c_\varepsilon(S)$.⁶

We will study in the next sections this specific mechanism, which we will abusively call pay-for-certainty from now on. The main feature of pay-for-certainty compared to the 0-correspondence elicitation is the additional payment for each chosen alternative. Observed choices are on the alternatives augmented by the payment, not on the original set of alternatives. The preference we want to elicit however is on the original set of alternatives. In the next section, we extend the set-up to take into account the payment and link it with the original set of alternatives.

A potential problem with pay-for-certainty with a uniform selection mechanism are individuals with a preference for randomization, as observed in Agranov and Ortoleva (2017) and modeled in Cerreia-Vioglio et al. (2018), for instance. Individuals with a preference for randomization will value the randomization process itself. Most of these models are built with risky or ambiguous lotteries in mind, which cannot be used with pay-for-certainty as presented here (Appendix B.2 explains why). More broadly, these models start with the idea that when outcomes are hard to distinguish, decision makers might value randomization over getting one for certainty, in particular in order to hedge between different risks. If the preference for randomization influences the preferences between outcomes in X , then we identify the order including the preference for randomization, and no problem arises. Now if x and y are hard to distinguish and it has some effects on the choice patterns, this will be identified with models of intransitive indifference or menu-dependence, and eliciting a choice correspondence is precisely the right thing to do here. One potential effect of preference for randomization is that indifferent decision maker will not select by themselves when $\varepsilon = 0$, and will choose the whole set. Again, this is not a problem for pay-for-certainty. It might solely imply that if the whole sample has a preference for randomization, eliciting the 0-correspondence is enough to identify indifference. It is an empirical question that will be tackled in Chapter 4

3.3 Identification of the Choice Correspondence

This section provides the formal conditions needed to identify the choice correspondence.

3.3.1 Assumptions

With pay-for-certainty with a uniform selection mechanism, decision makers choose a couple made of a set of alternatives and a bonus payment associated with the – size of – set. The natural

⁶ $c_\varepsilon(S)$ is the observed choice when the baseline payment is ε , see Definition 3.2.

choice space is the Cartesian product of the set of all non-empty subsets of X , and \mathbb{R} , call it Ω : $\Omega = \mathcal{P}(X) \times \mathbb{R}$. An element in this set is a couple (S, r) , where S is a non-empty subset of X , and r is a real. When r is positive, it is interpreted as a payment *to* the decision maker, when it is negative, as a payment *from* the decision maker. Call *set preferences* and note \succeq_2 the preferences of the decision maker on this new choice space.⁷ As usual, \sim_2 is the symmetric part of \succeq_2 , and \succ_2 is its asymmetric part.

It makes sense to relate set preferences with preferences, as Ω is partly built from $\mathcal{P}(X)$. First, we start with a remark on notation.

Remark (Notation when the real part is null). When the real part is equal to zero, that is, when we consider set preferences on the space $\mathcal{P}(X) \times \{0\}$, we omit the real part:

- For two sets S and S' in $\mathcal{P}(X)$, we note $S \succeq_2 S'$ for $(S, 0) \succeq_2 (S', 0)$;
- For two alternatives x and y in X , we note $x \succeq_2 y$ for $(\{x\}, 0) \succeq_2 (\{y\}, 0)$;
- We use similar abuses of notations for \succ_2 and \sim_2 .

With these notations in mind, we give the first links between preferences and set preferences.

Assumption 3.2 (Link between preferences and set preferences). *We impose some structure on set preferences \succeq_2 , in relations with properties of preferences \succeq . For any two elements $S, S' \in \mathcal{P}(X)$:*

1. *Preferences and set preferences are the same when payments associated with the alternatives are null, and the sets are singletons:*

$$\text{for all } x, y \in X, x \succeq y \Leftrightarrow (\{x\}, 0) \succeq_2 (\{y\}, 0)$$

2. *If for all elements $x \in S$ and for all elements $y \in S'$, $x \succeq y$, then $S \succeq_2 S'$. Additionally, if there exists $x \in S$ and $y \in S' \setminus S$ or $x \in S \setminus S'$ and $y \in S'$ with $x \succ y$, then $S \succ_2 S'$.*
3. *Take two alternatives $x, y \in X$ with $y \in S$, if $x \succ y$, then $S \cup \{x\} \setminus \{y\} \succ_2 S$.*
4. *Take two alternatives $x, y \in X$, if $x \notin S$ and there is no $y \in S$, $y \succ x$, then $S \cup \{x\} \succeq_2 S$.*

Property 1 identifies set preferences with preferences when sets are singletons, and the payments are null. Notice that we avoid as much as possible assumptions on \sim_2 , as its interpretation in our setup is not very clear. As we do not assume anything on \sim , we want to avoid as much as possible assumptions on \sim_2 that could be translated back by this property to \sim . Property 2 means that if all the alternatives in a set are at least as good as all the alternatives in another set, the first set is at least as good as the second one. For the set preference to be strict, the two sets must be different by at least one alternative, and this alternative must be strictly ordered with one in the

⁷It is for clarity of the exposition, as elements of X could be sets themselves, for instance, if we think of bundles of goods.

other set. Property 3 means that replacing an alternative in a set with an alternative that is strictly better imply that the new set is strictly better than the old one. Note that if x is in the original set, it implies that removing a strictly worse alternative from the original set yield a strictly better set, or the converse: adding a strictly worse alternative to the set yield a worse set. Property 4 means that adding a strictly undominated alternative yield a new set that is at least as good as the original one. These requirements are minor in the case of certain alternatives, and knowing that the decision maker will only get one of them, which is the focus of the pay-for-certainty procedure. They might not be so minor, however, in at least two cases: risky or uncertain lotteries, or if the decision maker gets the set at the end and the alternatives might be complement or substitute (say, a chair and a table).⁸

We impose some structure on set preferences, with three assumptions, *significant differences*, *quasi-linearity*, and *partial-transitivity*, and link preferences on X and preferences on Ω .

Assumption 3.3 (Monotonicity of set preferences). *If the only difference between two alternatives in Ω is the payment, the decision maker always prefers the highest payment to the lowest one.*

$$\text{for all } S \in \mathcal{P}(X), \text{ for all } r, r' \in \mathbb{R}, r > r' \Leftrightarrow (S, r) \succ_2 (S, r')$$

Assumption 3.4 (Significant differences on set preferences). *\succeq_2 is sensitive to small variations in the real part.*

$$\text{for all } S, S' \in \mathcal{P}(X) \text{ with } S \succeq_2 S' \text{ and } r, r' \in \mathbb{R} \text{ with } r > r', (S, r) \succ_2 (S', r')$$

Significant differences implies monotonicity and extends it to two sets that are comparable according to \succeq_2 .⁹ It is particularly relevant when two sets are indifferent. In that case, it tells us that a small payment in favor of one set will shift the preference towards this set. It does not impose anything on \sim , as the variations are on the real part, not on the set part.

Assumption 3.5 (Quasi-linearity of set preferences). *If one set is preferred to the other at a given level of payment, it is preferred at all levels of payments.*

$$\text{for all } S, S' \in \mathcal{P}(X), (S, 0) \succeq_2 (S', 0) \text{ if and only if, for all } r \in \mathbb{R}, (S, r) \succeq_2 (S', r)$$

Quasi-linearity expands the structure imposed in Assumption 3.2 from the real element being null to any value in \mathbb{R} , as long as it remains constant.

⁸Take two complementary events R (rain) and NR (not rain), and three (ambiguous) lotteries $x = (1, R; 0, NR)$, $y = (0, R; 1, NR)$ and $z = (1/3, R; 1/3, NR)$, which should be read as: with lottery x , the decision makers gets 1€ if it rains and 0 otherwise. A sufficiently ambiguity averse decision maker will prefer lottery z to lotteries x or y . He will, on the other hand, prefer lottery $(x + y)/2$ to lottery z . So he might prefer the set $\{x, y\}$ over the set $\{x, z\}$, which is a contradiction property 3.

⁹The name will be clear with Assumption 3.8.

Assumption 3.6 (Quasi-transitivity of set preferences). *For all S, S', S'' in $\mathcal{P}(X)$ and r, r', r'' in \mathbb{R} such that $(S, r) \succ_2 (S', r')$ and $(S', r') \succ_2 (S'', r'')$, then $(S, r) \succ_2 (S'', r'')$.*

Quasi-transitivity imposes transitivity of strict preferences, but nothing on indifference, i.e., on \sim_2 . Using identity, quasi-transitivity of set preferences implies quasi-transitivity of preferences, which is Assumption 3.1. The converse is not true in general.

Definition 3.2. (ε -correspondence). The ε -correspondence c_ε is the choice correspondence obtained on Ω when the bonus payment for choosing an alternative in S is equal to $\frac{1}{|S|}\varepsilon$.

The ε -correspondence is the observed choices. $c(S)$ is a theoretical object. We want to link c and c_ε . We use the classical idea that decision makers should choose a maximal alternative. Here, however, an alternative is not an element of X , but an element of Ω , as the alternative they have to choose from are from Ω .

Assumption 3.7 (Maximal Choice). *The decision maker chooses one of the maximal set according to \succeq_2 .*

$$c_\varepsilon(S) \in \operatorname{argmax}_{S' \subseteq S} \left\{ \left(S', \frac{|S'|}{|S|}\varepsilon \right) \right\}$$

This argmax is not necessarily unique, as the following example illustrates.

Example 3.2 (Uniqueness of the argmax). Take $S = \{x, y\}$ with preferences $x \sim y$ and no bonus payment ($\varepsilon = 0$). Then, using property 2 of Assumption 3.2, it is easy to see that $\{x\} \sim_2 \{x, y\} \sim_2 \{y\}$: the argmax is not unique.

3.3.2 Identification of the Choice Correspondence

We have talked so far of identifying the choice correspondence, that is, being able to *observe* $c(S)$. We provide now formal contents to this definition.

Definition 3.3 (Full Identification). We fully identify the choice correspondence of the decision maker when $c(S)$ is the only element in the $\operatorname{argmax}_{S' \subseteq S} \left\{ \left(S', \frac{|S'|}{|S|}\varepsilon \right) \right\}$ for a given ε .

This definition has two components. First, we must guarantee the uniqueness of the argmax. Second, this argmax must be the maximal alternatives in each set. We now give the conditions for these two components to be met.

3.3.2.1 Partial Identification of the Choice Correspondence

First, we need to introduce a preliminary result on the link between 0-correspondences, the choice correspondence, and the ε -correspondence when $\varepsilon > 0$, i.e., between c_0 , c , and c_ε . The

0 -correspondence is the observed choices when there is no incentive to choose several alternatives. The ε -correspondence when $\varepsilon > 0$ is the observed choices when there is a strict positive incentive to choose several alternatives. Formally, for all sets $S \in \mathcal{P}(X)$, $c_0(S)$ is one element of $\operatorname{argmax}_{S' \subseteq S} \{(S', 0)\}$ and $c_\varepsilon(S)$ is one element of the $\operatorname{argmax}_{S' \subseteq S} \left\{ \left(S', \frac{|S'|}{|S|} \varepsilon \right) \right\}$. Before looking at the full identification of choice correspondences, let us introduce a preliminary *partial* identification result, which will help us understanding when the full identification happens.

Proposition 3.1 (Partial Identification). *When strong monotonicity, the structure imposed in Assumption 3.2 and maximal choice are satisfied, for all $\varepsilon > 0$, $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$. It is true for c_0 and any c_ε .*

Proof. Take $S \in \mathcal{P}(X)$. We prove the first inclusion and then the second inclusion by contradiction. First inclusion, $c_0(S) \subseteq c(S)$:

- Imagine it is no the case. There exists an alternative x in $c_0(S)$ which is not in $c(S)$.
- $c(S)$ is made of all the undominated alternatives in S so that x must be dominated by another alternative y in S ($y \succ x$).
- Take the set $S' = c_0(S) \cup \{y\} \setminus \{x\}$. x does not belong to $c_0(S)$, we have $S' \succ_2 c_0(S)$ (by property 2 from Assumption 3.2). This is in contradiction with the fact that $c_0(S)$ is in the argmax when $\varepsilon = 0$.
- Therefore, $c_0(S)$ must contain only maximal alternatives in S and $c_0(S) \subseteq c(S)$.
- Example 3.2 above makes clear why it might not be an equality.

Second inclusion, $c(S) \subseteq c_\varepsilon(S)$:

- Imagine it is not the case. There exists an alternative x in $c(S)$ which is in not in $c_\varepsilon(S)$.
- Take $S' = c_\varepsilon(S) \cup \{x\}$. x is a maximal alternative, so by property 4 in Assumption 3.2, we have $S' \succeq_2 c_\varepsilon(S)$.
- x is not in $c_\varepsilon(S)$, so that $|S'| - 1 = |c_\varepsilon(S)|$.
- By strong monotonicity, we therefore have that $\left(S', \frac{|S'|}{|S|} \varepsilon \right) \succ_2 \left(c_\varepsilon(S), \frac{|c_\varepsilon(S)| - 1}{|S|} \varepsilon \right)$ Therefore $\left(c_\varepsilon(S), \frac{|c_\varepsilon(S)|}{|S|} \varepsilon \right)$ is strictly dominated when the additional payment is $\frac{1}{|S|} \varepsilon$ which is a contradiction with the fact that it is an element of the argmax .
- It implies that $c(S) \subseteq c_\varepsilon(S)$.

□

Proposition 3.1 is easy to check in practice, as long as we observe the 0-correspondence and one ε -correspondence. Chapter 4 gives an illustration of how to do it.

3.3.2.2 Full Identification of the Choice Correspondence

Proposition 3.1 does not guarantee that we can fully identify the choice correspondence of the decision maker. It only guarantees the partial identification of the choice correspondence. There is an obvious corollary, however.

Corollary 3.1 (Full Identification by Set Inclusion). *When for one $\varepsilon > 0$ and for all S in $\mathcal{P}(X)$, we have that $c_0(S) = c_\varepsilon(S)$, we fully identify the choice correspondence of the decision maker. Moreover, $c(S) = c_0(S) = c_\varepsilon(S)$.*

Corollary 3.1 tells us that if we are lucky enough, it is possible to identify the choice correspondence of the decision makers. One difficulty of this result lies in the 0-correspondence. As pointed out in Example 3.2, the argmax may not be unique, and in particular, it may not be exactly the set of all maximal alternatives. One last result gives us a hint on when full identification is more likely to happen. We have to introduce one last assumption in order to do so, *insignificant difference*.

Assumption 3.8 (Insignificant Difference). \succ_2 is not sensitive to small variations in \mathbb{R} .

For all $S, S' \in \mathcal{P}(X)$, and for all $r, r' \in \mathbb{R}$ with $(S, r) \succ_2 (S', r')$, there exists a $t > 0$ such that for all t' with $0 \leq t' \leq t$, we have $(S, r) \succ_2 (S', r' + t')$.

In other words, if a couple (S, r) is strictly preferred to another one (S', r') , a small enough payment will not reverse the preference. It is akin to a continuity assumption of set preferences, but only on \mathbb{R} . It is a kind of counterpart to the significant difference assumption introduced earlier, which introduces a dissymmetry between \succ_2 and \sim_2

Proposition 3.2 (Full Identification). *Under significant and insignificant difference, quasi-linearity, and quasi-transitivity, there exists $\varepsilon > 0$ such that for all ε' with $0 < \varepsilon' < \varepsilon$, the argmax is unique and equal to c : $c_{\varepsilon'} = c$.*

Proof. Take a set $S \in \mathcal{P}(X)$. The case where $c(S) = S$ is trivially true. The interesting cases are when $c(S) \subset S$.

- First, notice that by the results of the Proposition 3.1, for any $\varepsilon > 0$, the only sets that can be chosen are supersets of $c(S)$.
- Take any strict superset of $c(S)$, call it S' . By definition, $c(S) \succ_2 S'$.
- By insignificant difference, there exists $t_{S'}$ such that $(c(S), 0) \succ_2 (S', t_{S'})$
- Take now any $\varepsilon \geq 0$, by quasi-linearity, we have that $(c(S), \frac{|c(S)|}{|S|}\varepsilon) \succ_2 (S', \frac{|c(S)|}{|S|}\varepsilon + t_{S'})$.
- Take now $0 < \varepsilon' < \min_{\{S' | c(S) \subset S' \subseteq S\}} \frac{|S|}{|S'| - |c(S)|} t_{S'}$. This minimum exists as X , and, therefore, S is finite. It is well defined, as $|S'| - |c(S)| \geq 1$.
- For that ε' and for all S' strict superset of $c(S)$, $(S', \frac{|c(S)|}{|S|}\varepsilon' + t_{S'}) \succ_2 (S', \frac{|S'|}{|S|}\varepsilon')$. It implies, by quasi-transitivity, that $(c(S), \frac{|c(S)|}{|S|}\varepsilon') \succ_2 (S', \frac{|S'|}{|S|}\varepsilon')$.

- So for all ε' define as above, only $c(S)$ is in the argmax.

□

Proposition 3.2 tells us that we should use the smallest positive ε possible. In practice, however, we are bounded below on the value of ε , and therefore might not be able to reach a sufficiently small one. We might have to rely on partial identification only.

3.3.3 Consistency when Full Identification Fails

It is a problem from a theoretical perspective. If a 0-correspondence and an ε -correspondence yield partial identification, but none satisfy a property, it does not mean that the choice correspondence of the decision maker, which is in between, would not satisfy it. We tackle this question, by looking under which conditions a property could be satisfied in practice when there is partial identification but not full identification.

Definition 3.4 (Compatibility with a Property). A pair (c_0, c_ε) with $\varepsilon > 0$ is said to be *compatible with a property P* if:

1. For all S in $\mathcal{P}(X)$, $c_0(S) \subseteq c_\varepsilon(S)$;
2. There exists a choice correspondence c , for all $S \in \mathcal{P}(X)$, $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ and c satisfies property P .

Compatibility with a property means that we cannot reject the fact that the decision maker satisfies that property. We may not observe its satisfaction because our tools are too limited to observe it, not because the decision makers do not satisfy it. On the other hand, the decision maker is not compatible with a property implies that the decision maker cannot satisfy it. An obvious sufficient condition for compatibility with property P is when c_0 or c_ε satisfy P . As we will show soon, it is not the only possible cases, however.

Unfortunately, there is no general method to tell if there exists a choice correspondence between a 0-correspondence and an ε -correspondence that satisfies a property when neither satisfies it. One possibility in practice is to check the property for all possible choice correspondences between the 0-correspondence and the ε -correspondence. This brute force method might quickly be computationally heavy, however.¹⁰ In the following, we provide conditions for compatibility with classical preferences, intransitive indifference models, and the menu and context-dependent models, that do not rely on the brute force method.

¹⁰We will use it for some properties in Chapter 4.

3.3.3.1 Compatibility with Classical Preferences

First, let us provide an example where a pair (c_0, c_ε) is compatible with classical preferences.

Example 3.3 (Compatibility with Classical Preferences). Take $X = \{x, y, z\}$, and the observed choices in Table 3.2. The pair (c_0, c_ε) satisfies the first condition of compatibility with classical preferences. What about the second condition? Note that classical preferences rationalize neither c_0 nor c_ε .

Table 3.2: The observed choices of one individual, with and without the ε bonus payment.

Choice Set	0-correspondence	ε -correspondence	Choice correspondence
$\{x, y\}$	$\{x\}$	$\{x, y\}$	$\{x\}$
$\{x, z\}$	$\{x, z\}$	$\{x, z\}$	$\{x, z\}$
$\{y, z\}$	$\{z\}$	$\{z\}$	$\{z\}$
$\{x, y, z\}$	$\{x\}$	$\{x, y, z\}$	$\{x, z\}$

It is possible however to build a choice correspondence c between c_0 and c_ε that would be rationalized by a classical preference, as the last column of Table 3.2 shows. For every set S , $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ and c is rationalized by classical preferences. In fact, the revealed classical preference is $x \sim z \succ y$. In this case, it is the unique preference relation which is compatible with both the 0 and ε -correspondence.

Generally speaking, compatibility with classical preferences is more natural to tackle from a preference standpoint, rather than looking for a choice correspondence that satisfies WARP. The compatibility of a pair (c_0, c_ε) with classical preference imposes some conditions on the strict preferences \succ and the indifference \sim .

1. On the strict preferences, it is clear that if an alternative is not chosen with a bonus, then surely it is worse than at least one chosen alternative. For all unchosen alternatives $y \in S \setminus c_\varepsilon(S)$, there must exist an $x \in c_\varepsilon(S)$, $x \succ y$. It does not have to be all alternatives that are not chosen are worse than all alternatives that are chosen, as potentially only some of the chosen alternatives (i.e., in $c_\varepsilon(S)$) will be in $c(S)$.
2. It is also the case if we want $x \succ y$, then y can never be chosen when x is available. Formally, it implies that if there exists S , with x and y in S , and $y \in c_0(S)$, then not $y \succ x$. Moreover, we might suspect that if for all S , with $x, y \in S$, y is never chosen, then $x \succ y$.
3. On the indifference part, it is also clear that if two alternatives are chosen together with no bonus payments, they are indifferent. Formally, for two alternatives x, y , if there exists S such that $x, y \in c_0(S)$, then certainly x and y are indifferent ($x \sim y$). Indeed, $x, y \in c(S)$.

Definition 3.5 (Compatible Classical Preference). We propose a constructive method to build a classical preference \succeq which is compatible with the pair (c_0, c_ε) . For all $x, y \in X$:

- $x \succ y$ if and only if both
 1. For all S , $x, y \in S$ and $y \notin c_0(S)$ (which implies that $c_0(\{x, y\}) = \{x\}$)
 2. There is no sequence of integer $i = 1, \dots, n$ such that $y \in c_0(S_1), x \in S_n$ and $S_i \cap c_0(S_{i+1}) \neq \emptyset$ (an acyclicity condition). Note that this condition implies the first one when $n = 1$.
- $x \sim y$ otherwise.

Note that this defines a unique compatible classical preference for each pair (c_0, c_ε) . We have not proved yet that this preference is classical, which is the object of the next proposition.

Proposition 3.3. *The compatible classical preference is reflexive, transitive, and complete, i.e., it is a classical preference.*

Proof. Reflexivity does not have to be proven.

Completeness is evident from the construction.

Transitivity remains. Take three alternatives x, y , and z such that $x \succ y \succ z$. Is it the case that $x \succ z$? There are four possibilities:

1. $x \succ y \succ z$.
2. $x \sim y \sim z$
3. $x \succ y \sim z$
4. $x \sim y \succ z$

Take the first case. We want to show that $x \succ z$:

1. Imagine it is not the case. There exists a sequence S_1, \dots, S_n of subsets of X , $z \in c_0(S_1), x \in S_n$ and $S_i \cap c_0(S_{i+1}) \neq \emptyset$ (possibly with $n = 1$).
2. It contradicts $x \succ y$. Indeed, $y \succ z$ implies that $y \in c_0(\{y, z\})$. We have that $z \in \{y, z\} \cap c_0(S_1)$, which implies that the intersection is not empty. And finally, $x \in S_n$. We have used the sequence between z and x and added a first set $S_0 = \{y, z\}$ to build a sequence between y and x . However, this is impossible because of $x \succ y$ – so $x \succ z$.

Note that the proof works identically if $y \sim z$, so we have proven cases 1 and 3. It is also similar for case 4.

Let us prove case 2:

1. $x \sim y$ means that there exists a sequence S_1, \dots, S_n of subsets of X , $y \in c_0(S_1)$, $x \in S_n$ and $S_i \cap c_0(S_{i+1}) \neq \emptyset$ (possibly with $n = 1$). $y \sim z$ means that there exists a sequence S'_1, \dots, S'_m of subsets of X , $z \in c_0(S'_1)$, $y \in S'_m$ and $S'_i \cap c_0(S'_{i+1}) \neq \emptyset$ (possibly with $m = 1$).
2. Build a new sequence between x and z by noting that $y \in S'_m \cap c_0(S_1)$, which implies that the intersection is not empty and we can concatenate the sequences.

□

Definition 3.6 (Associated Choice Correspondence). Take \succeq a preference which is reflexive and quasi-transitive. Build the *choice correspondence associated with \succeq , c^\succeq* , by:

$$\text{For all } S \in \mathcal{P}(X), c^\succeq(S) = \{x \in S \mid \text{there is no } y \in S, y \succ x\}$$

When preferences are classical, we also have:

$$\text{For all } S \in \mathcal{P}(X), c^\succeq(S) = \{x \in S \mid \text{for all } y \in S, x \succeq y\}$$

This is a consequence of the definition of the associated choice correspondence when preferences are transitive and complete. The associated choice correspondence represents the theoretical choice correspondence of the preferences built. As the compatible preference classical preference is a classical preference, the associated choice correspondence satisfies WARP.

Proposition 3.4 (Compatibility with Classical Preferences). $c_0(S) \subseteq c^\succeq(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$ and c^\succeq satisfies WARP if and only if there exists a choice correspondence c , $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$ and c satisfies WARP.

Proof. The if direction is obvious, we have to prove the only if part.

Define \succeq^c , the revealed preference of c , according to Definition 1.5. Because c satisfies WARP, \succeq^c is transitive and complete. If $x \succ^c y$, then for all S with $x, y \in S$, $y \notin c(S)$, which implies $y \notin c_0(S)$. It implies that $x \succ y$. Indeed, if it were not the case, then the acyclicity condition for c_0 would be violated, which means that it would be violated for c , and \succeq^c could not be transitive. So, if $x \succ^c y$, $x \succ y$. So $\succ \subseteq \succeq^c$. So $c^\succ(S) \subseteq c(S) \subseteq c_\varepsilon(S)$.

Let us show now that $c_0(S) \subseteq c^\succeq(S)$. Imagine it is not the case. Then there exists $x \in c_0(S)$, $x \notin c^\succ(S)$. $x \notin c^\succ(S)$ means that there exists $y \in S$, $y \succ x$, which implies that $x \notin c_0(S)$, which is a contradiction. So we have proved that $c_0(S) \subseteq c^\succeq(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$. □

Corollary 3.2 (Minimality of the Indifference). *The compatible classical preference is the classical preference compatible with the pair (c_0, c_ε) with the least indifference.*

Proof. In the proof of Proposition 3.4. □

In practice, to check the compatibility with classical preferences, it is easier to build the compatible preference using only the first part of the definition of \succ , and then check whether it is acyclic and whether $c^\succeq(S)$ is included in $c_\varepsilon(S)$ for all S . The algorithm proposed yield the compatible classical preference with the least indifference. This compatible preference is not unique, however, as Example 3.4 shows.

Example 3.4 (c is not unique). Take $X = \{x, y, z, t\}$, and the choices in Table 3.3

Table 3.3: A pair (c_0, c_ε) where the compatible preference is not unique.

Choice Set	c_0	c_ε
$\{x, y\}$	$\{x, y\}$	$\{x, y\}$
$\{x, z\}$	$\{x\}$	$\{x, z\}$
$\{y, z\}$	$\{z\}$	$\{y, z\}$
$\{x, y, z\}$	$\{x\}$	$\{x, y, z\}$
$\{x, t\}$	$\{x\}$	$\{x\}$
$\{y, t\}$	$\{y\}$	$\{y\}$
$\{z, t\}$	$\{z\}$	$\{z\}$
$\{x, y, t\}$	$\{x\}$	$\{x, y\}$
$\{x, z, t\}$	$\{x\}$	$\{x, z\}$
$\{y, z, t\}$	$\{y\}$	$\{y, z\}$
$\{x, y, z, t\}$	$\{x\}$	$\{x, y, z, t\}$

The compatible preference is $x \sim y \succ z \succ t$. It is not unique, as another preference is compatible: $x \sim y \sim z \succ t$.

3.3.3.2 Compatibility with a Partial Order

In Section 1.2.1, we have given three intransitive indifference models. Partial orders is the largest class of intransitive indifference preferences, and contain semi-orders and interval orders. This particular intransitive indifference model is observationally equivalent to incomplete preferences, where no relation would replace the indifference part of the relation. For these two reasons, we will focus solely on the study of partial orders with partial identification, and not explore semi-orders and interval orders.

It is again easier to tackle the study of compatibility with partial orders directly with preferences, rather than trying to find a choice correspondence which would be compatible with the pair (c_0, c_ε) and would satisfy WARNI. The conditions imposed on preferences are the same as with classical preferences. The difference between classical preferences and partial orders is on the indifference

part: it does not have to be transitive for partial orders. It motivates the definition of a compatible partial order.

Definition 3.7 (Compatible Partial Order). We propose a constructive method to build a partial order \succeq which is compatible with the pair (c_0, c_ε) . We have three steps:

1. $x \succ y$ if:
 1. There exists $S \in \mathcal{P}(X)$, $x \in c_\varepsilon(S)$ and $y \in S \setminus c_\varepsilon(S)$;
 2. For all $S \in \mathcal{P}(X)$, if $x, y \in S$, $y \notin c_0(S)$.
2. Close \succ transitively;
3. Build $x \sim y$ when not $x \succ y$ and not $y \succ x$.

Because we close transitively \succ , it might not be a partial order, as there might be a cycle created. It is not the case, and this is the purpose of the next proposition.

Proposition 3.5. *The compatible partial order is a partial order.*

Proof. The compatible partial order is transitive by definition. We have to prove the absence of cycles in the partial order. Assume it would be the case. That is, there is a sequence $x_1, \dots, x_i, \dots, x_n$ such that $x_i \succ x_{i+1}$ and $x_n \succ x_1$. The transitive closure cannot generate cycles if it were not originally here. So we can assume that the cycles were cycles generated from the observations. Take the set made of all the alternatives in the cycle: $S = \{x_1, \dots, x_n\}$. We can do that because we have assumed full observability. $c_0(S)$ is non-empty, implying that one alternative at least is chosen in S . Call x_k this alternative. We cannot have that $x_{k-1} \succ x_k$, because of condition 2 in the definition of a compatible partial order, so we have a contradiction and \succ is acyclic. \square

We use the choice correspondence associated with the partial order as defined in Definition 3.6. Compared to the previous subsection, we can only use the first and not the second definition of the associated choice correspondence. As the compatible partial order is a partial order, the associated choice correspondence satisfies WARNI.

Proposition 3.6 (Compatibility with a Partial Order). $c_0(S) \subseteq c^\succeq(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$ and c^\succeq satisfies WARNI if and only if there exists a choice correspondence c , $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$ and c satisfies WARNI.

Proof. The if part is obvious, we have to prove the only if part.

Define \succ^c , the revealed preference of c , according to Definition 1.5. Because c satisfies WARNI, \succ^c is transitive. We have two cases.

1. Either there exists a set $S \in \mathcal{P}(X)$, such that there exists $y \in S \setminus c_\varepsilon(S)$.

2. Or for all $S \in \mathcal{P}(X)$, $c_\varepsilon(S) = S$.

Case 2 is trivial because everything is chosen all the time with c_ε , \succ is empty, and c^\succ satisfies WARP and therefore WARNI.

Case 1 is the interesting case. It implies that $y \in S \setminus c(S)$. Therefore there exists $x \in c(S)$, $x \succ^c y$ (remember that \succ^c is transitive). Because $x \succ^c y$, for all S such that $x, y \in S$, $y \notin c(S)$. It implies that for all S such that $x, y \in S$, $y \notin c_0(S)$, which implies that $x \succ y$. It implies that for any pair such that $x \succ y$, we must have $x \succ^c y$ (both are transitive). So $\succ \subseteq \succ^c$, which means that for all $S \in \mathcal{P}(X)$, $c(S) \subseteq c^\succ(S)$. It proves that $c_0(S) \subseteq c^\succ(S)$ for all $S \in \mathcal{P}(X)$.

The inclusion $c^\succ(S) \subseteq c_\varepsilon(S)$ comes from the repetition of the reasoning above. As for any $y \in S \setminus c_\varepsilon(S)$, there exist an $x \in c(S)$, $x \succ^c y$, we will have that $x \succ y$, and therefore $y \in S \setminus c^\succ(S)$. We have that $S \setminus c^\succ(S) \subseteq S \setminus c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$, which is equivalent to $c^\succ(S) \subseteq c_\varepsilon(S)$ for all $S \in \mathcal{P}(X)$. \square

Corollary 3.3 (Minimality of the Strict Preference). *The compatible partial order is the partial order compatible with (c_0, c_ε) with the least strict preferences.*

Proof. In the proof of Proposition 3.6, we have shown that $\succ \subseteq \succ^c$, which means that any partial order compatible with (c_0, c_ε) has at least as many strict preferences as the compatible partial order. \square

Proposition 3.6 guides us on how to build a partial order compatible with any pair (c_0, c_ε) . It also tells us when there is no compatible partial order. In practice, the procedure is to build the compatible order and the associated choice correspondence, and then to check that for all $S \in \mathcal{P}(X)$, we have $c_0(S) \subseteq c^\succ(S) \subseteq c(S)$. In that case, the pair (c_0, c_ε) is compatible with a partial order. Otherwise, it is not. Corollary 3.3 shows that the compatible partial order is the one with the least strict relation. In other words, it provides the least information possible on the preference of the decision maker. If we want to be conservative from a welfare standpoint, it is the right preference to use. Note that because a choice correspondence which satisfies WARP also satisfies WARNI, compatibility with classical preferences implies compatibility with a partial order. It means that Example 3.4 shows that the compatible partial order is not unique.

3.3.3.3 Compatibility with Menu-Dependent Threshold

As given in Section 1.2.1.5, a necessary and sufficient condition for a choice correspondence to be rationalized by the menu-dependent threshold is to satisfy functional acyclicity. Functional acyclicity is satisfied if and only if there are no cycles of strict preferences. A necessary condition for the compatibility of a pair (c_0, c_ε) with the menu-dependent threshold model mirrors this condition.

Definition 3.8 (Strict Cycle of Observations with Partial Identification). A *strict cycle of observation* are n sets S_1, \dots, S_n in $\mathcal{P}(X)$ such that:

$$\begin{aligned} S_1 \setminus c_\varepsilon(S_1) \cap c_0(S_2) &\neq \emptyset \\ S_2 \setminus c_\varepsilon(S_2) \cap c_0(S_3) &\neq \emptyset \\ &\dots \quad \dots \quad \dots \\ S_n \setminus c_\varepsilon(S_n) \cap c_0(S_1) &\neq \emptyset \end{aligned}$$

Proposition 3.7 (Compatibility with Functional Acyclicity). A *necessary condition for a pair (c_0, c_ε) to be compatible with functional acyclicity is to not have a strict cycle of observation with partial identification.*

Proof. By definition of partial identification, for all S in $\mathcal{P}(X)$, $c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S)$. It means that for all S , $S \setminus c_\varepsilon(S) \subseteq S \setminus c(S)$. If a strict cycle of observations with partial identification exists, it implies that any choice correspondence will have a strict cycle of observations, and thus will violate functional acyclicity. \square

3.3.3.4 Compatibility with Fixed Point

As given in Section 1.2.1.6, a necessary and sufficient condition for a choice correspondence to be rationalized by the context-dependent threshold model is to satisfy fixed point (see Axiom 1.9). Fixed point requires the existence of *one* alternative that is chosen in a superset, and in any subset, the condition can be directly imposed on c_ε .

Proposition 3.8 (Compatibility Fixed Point). A *pair (c_0, c_ε) is compatible with fixed point if and only if c_ε satisfies fixed point.*

In practice, it means that as long as c_ε satisfies fixed point, we can find a c that satisfies it.

3.4 Conclusion

In this chapter, we have discussed the difference between the empirical and theoretical literature on revealed preferences. Most of the theoretical literature, and in particular the literature interested in the relaxation of the classical paradigm, start with a choice correspondence, whereas most of the empirical literature identifies a choice function. Decision makers are assumed to choose all the maximal alternatives in a given choice set, whereas, in practice, they are forced to choose a single alternative in a given choice set.

Incentivizing decision makers to choose exactly their maximal alternatives is not easy in practice. We introduced a new method to do so, *pay-for-certainty with a uniform selection mechanism*.

We provide two conditions under which the choice correspondence is identified. One is testable in practice, but restrictive, and another that is likely to be more general, but not testable. Full identification might be hard to get in practice, so we also characterize the compatibility with different properties when we have a weaker partial identification.

Chapter 4

Pay-for-certainty in an Experiment

4.1 Introduction

This chapter is dedicated to an experiment we ran using pay-for-certainty with the uniform selection mechanism. It has two main objectives.

First, it illustrates the method introduced in Chapter 3, and particularly the central characterization introduced in Section 3.2.3: pay-for-certainty with a uniform selection mechanism.

From the choice set S , each alternative chosen adds a *bonus payment* of $\frac{1}{|S|}\varepsilon > 0$ per alternative to the gain of the subject. The total bonus payment is $\frac{|c_\varepsilon(S)|}{|S|}\varepsilon$ in each choice set. The alternative he gets is selected using a uniform random draw on the set of chosen alternatives $c_\varepsilon(S)$.¹

We run the experiments with three main characterizations and leverage the identification results of Chapter 3. First, we identify a choice function using the classical forced-choice experimental method. Subjects have to choose precisely *one* alternative in each choice set. We call this method *forced single-choice*. Second, pay-for-certainty with a null bonus payment called the *0-correspondence* hereafter. It is the simplest weakening of the classical method allowing for multiple choices. Third, pay-for-certainty with a total bonus payment of ε of 1 cent, called the *1-correspondence* hereafter. We then use the limit and set-inclusion identification results proven in Chapter 3. The latter is more complicated to implement than the former, as it relies on both the 0-correspondence and on the 1-correspondence for each subject, whereas the former only relies on the 1-correspondence. In exchange, the latter guarantees that we identify the choice correspondence of the subjects.

Second, the experiment aims at getting new insights into individual decision making. Running an experiment where menus are chosen, rather than single alternatives, allows us to explore models beyond classical preferences, and in particular models of intransitive indifference and menu-dependent

¹Remember that $c_\varepsilon(S)$ is the set of all chosen alternatives in S , when the bonus payment is $\frac{1}{|S|}\varepsilon > 0$.

choice presented in Section 1.2. Excluding approval voting experiments, very few incentivized experiments have allowed decision makers to choose several alternatives. To the best of our knowledge, five have done that so far: Danan and Ziegelmeyer (2006), Costa-Gomes, Cueva, and Gerasimou (2016), “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019), Ong and Qiu (2018) and Agranov and Ortoleva (2017). It is at odds with many decision-theoretic models which require menu choice, e.g., Arrow (1959), Amartya Sen (1971), Schwartz (1976), Eliaz and Efe A. Ok (2006), Aleskerov, Bouyssou, and Monjardet (2007).

Another benefit of menu choices is to enable the study indifference directly. Choice functions reveal indifference directly we use weak revealed preferences, which yield to a lower bound on it. Choice correspondences, on the other hand, accurately quantify it. It might have important individual and collective welfare implications. Arrow’s impossibility theorem shows that it is impossible to aggregate the preferences of individuals in a non-dictatorial manner. It is evident, however, that in the extreme case where all decision makers are indifference between all alternatives, there will be no impossibility theorem to aggregate their preferences. Maniquet and Mongin (2015) have shown a stronger result: when preferences are can be divided into two indifference classes, there is no impossibility to aggregate preferences either.

In the experiment, subjects chose between four different paid tasks: an addition task, a spellcheck task, a memory task, and a copy task. The use of tasks in experimental economics is not new. For instance, Augenblick, Niederle, and Sprenger (2015) used effort tasks to measure time-inconsistency. Tasks have several advantages for this experiment. First, these tasks were built for this experiment, and thus, subjects did not know anything about them before the session. It allows us to control their knowledge about the tasks. There were three levels of information, a very shallow one, one where they could see how the task looked like, and finally one where they trained on the task before any choice. Second, preferences are more likely to be sharp with effort tasks than between lotteries, as shown by Augenblick, Niederle, and Sprenger (2015), e.g., some subjects might hate calculus and thus will actively reject the addition task. We want to create an environment where indifference is not likely to be widespread. Finally, there is no objective ranking of the tasks, and they are quite different. It should yield different preferences, which is of interest to study preference aggregation.

In order to fully falsify the different models we test, all subjects faced choices in all possible subsets of the four incentivized tasks and thus faced at least 11 choices. According to the theory of Chapter 3, we need to observe choices with no gain (0 cent) and choices with positive, and small gains compared to the payoffs of the alternatives ($\varepsilon > 0$). We went a bit further, as most subjects faced *three* different levels of bonus payments, i.e., three different ε : a *no* (0 cent), a *low* (1 cent), and a *high* (12 cents) gain levels. It helps us characterize the validity of the limit identification result proven in Proposition 3.2. Crucially, the maximal bonus payment for adding alternative was never more than 6 cents, which is lower than the lowest increment for additional success in any task (10 cents). We also forced some subjects to choose a single alternative. For some of these subjects, we

have both a forced single choice elicitation and 0- and 1-correspondences. Results with high gains are shown in Appendix C.5. They are not significantly different from the results with low gains, so only results with no and low gain will be presented in the core of the dissertation.

We show the results of 214 participants in various specifications of the experiment, spanning across 13 sessions.² Overall, subjects took the opportunity to choose several alternatives. In the 0-correspondence, unincentivized to choose several alternatives, only 55.40% of all choices made are single alternatives, and only 15% of subjects always chose single alternatives. Using the characterization of Proposition 3.1 and Corollary 3.1, we fully identify the choice correspondence of 18% of the subjects, and we partially identify it for another 40%.

Classical preferences rationalize 57% of forced single choice, the benchmark experimental method. Under the assumption that 1 cent is low enough to use the limit identification of Proposition 3.2, 40% of choice correspondences can be rationalized by classical preferences. Classical preferences rationalize 97% of fully identified choice correspondences and are compatible with 92% of partially identified choice correspondences. The difference between choice functions and identified choice correspondences is always significant.³

Classical preferences rationalize or are compatible with most fully or partially identified choice correspondences, but not all 1-correspondences. In the latter case, we go beyond classical preferences and look at intransitive indifference models and menu-dependent models introduced in Section 1.2. Intransitive indifference models rationalize only marginally more 1-correspondences, i.e., 42%. The introduction of menu-dependence raises the rationalizability of the sample. The occasional optimality model rationalizes 57% of 1-correspondences, the menu-dependent model 91% and the context-dependent one 96%, all significantly higher than classical preferences. [Using a Fisher exact test.]

We then explore what kind of preferences subjects express when their preferences are classical. We use strict revealed preferences in the central assessment of forced single choice; it implies that indifference is not elicited. There are four tasks in the experiment. Therefore a complete preference relation is made of 6 binary relations. On the one hand, with forced single choices, these six binary relations are strict preferences. On the other hand, when we fully identify the choice correspondences, the number of binary relations that are indifference is for each subject is on average 2.12.⁴ When we use the partially identify choice correspondence, the figure is 2.08. With 1-correspondences, the figure is 3.11. In all cases when subjects can choose several alternatives, subjects express significant indifference that is not captured by choice functions.⁵

²We dropped some subjects; the restrictions are given in Section 4.2.6.

³The p-value of the two-sided two-sample unequal variance t-test is of 0.04 for the test between 1-correspondences and choices functions and below 0.001 with partial and full identification.

⁴We have excluded subjects who chose all the time all tasks, on whom the experiment has no bite.

⁵This result is robust to using weak revealed preferences with choice functions, as shown in Appendix C.4. With that interpretation, the number of indifference relations with choice function is 0.29.

Finally, we show that when subjects have more information, their choice is more precise. That is the size of their chosen sets decreases, as well as their number of indifference relations. Indeed, we might think that if decision makers have little information about two tasks, it is hard for them to value the tasks. In that case, they might sometimes choose one, sometimes the other.⁶ This is precisely the kind of behaviors the models of menu and context-dependent choice of Aleskerov, Bouyssou, and Monjardet (2007), Frick (2016), and Tyson (2018) can rationalize.

We first start this chapter by describing experiments with menu-choices in Section 4.1.1 Section 4.2 explains the experimental design and describes the sample. Section 4.3 provides the results with classical preferences. Section 4.4 takes advantage of the elicitation of menu choices to go beyond classical preferences. Section 4.5 explores the influence of the information provided to subjects on the results. Finally, Section 4.6 concludes.

4.1.1 Menu Choice in the Literature

Five experiments have tried to elicit menu choice in practice, mainly in attempts to explore incomplete preferences. In the first one, Danan and Zieglmeyer (2006) looked at the choice between a lottery and certain amounts using a bracketing procedure. Decision makers could postpone the choice, at a cost. This design essentially implements the *flexibility* selection mechanism given in Section 3.2.2.2, in the context of the choice of lotteries. They found a significant proportion of decision makers exhibiting incomplete preferences, that is, preferred to postpone the choices between one lottery and one certainty equivalent.

Costa-Gomes, Cueva, and Gerasimou (2016) use a non-forced-choice procedure in the choice between headphones. Subjects could postpone the choice, at a cost, which them to have a look at the headphones. It provided them with more information about choice objects. They also find a significant proportion of decision makers who postponed, and some subjects were closer to being rationally indecisive and 73% looked like classical decision makers in a non-forced choice setting.

“Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019) and Ong and Qiu (2018) both allowed decision makers to delegate their choices to a random device. The name might suggest it is an instance of the *delegation selection mechanism*, but it is close to the *uniform selection mechanism*. “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019) implement an experiment testing the completeness axiom with ambiguous prospects, looking for Bewley (2002) preferences. They perform three experiments where decision makers had to choose between risky and ambiguous lotteries. The key is, decision makers were allowed to delegate their choice to a random device. A significant proportion of decision makers did so. Ong and Qiu (2018) built an experiment around the ultimatum game. Proposers faced binary choices between the equal allocation and (random) series of unequal

⁶It is related to, but different from, the notion of preference imprecision (see Bayrak and Hey (2019) for a review).

allocation. Receivers face binary choices between accepting and rejecting the proposed allocations. Both receivers and proposers can randomize their choices, and had to state a willingness-to-pay for the randomization. Randomizing is only understandable if subjects have incomplete preferences, as is a positive willingness-to-pay. They found that many subjects randomize at some point, a strong indication that subjects value having the possibility to choose several alternatives, and potentially have incomplete preferences.

Finally, Agranov and Ortoleva (2017) ran an experiment where decision makers faced several times the same choices between lotteries. Crucially, even though they were aware that the choices were the same decision makers still changed their choices in some specifications. It indicates that they were not so sure about what the best alternative was.

Except for Agranov and Ortoleva (2017), it is costly to choose several alternatives in these experiments. It is, therefore, a dominant strategy for subjects indifferent between two or more alternatives to select only one. In Agranov and Ortoleva (2017)'s experiment, it is not a strictly dominant strategy to switch between the repetition of the same choice when indifferent. All these experiments fail or might fail to estimate the extent of the indifference of subjects adequately. Another difference, except for the experiment run by Costa-Gomes, Cueva, and Gerasimou (2016), is that all the experiments used the choice between risk or ambiguous lotteries, and thus differed in a crucial way from the choice of tasks implemented here. The main differences between the experiment of Costa-Gomes, Cueva, and Gerasimou (2016) are the selection mechanism, and the cost of not choosing, which is much higher in their experiment compared to here.

4.2 Design of the Experiment

Subjects chose between four different incentivized tasks. They chose three times in all possible subsets of alternatives. Each time with a different gain level or forced single choice. For subjects who chose according to the forced single choice procedure, they always performed it first. We feared that because the forced single choice made subjects think of one alternative in each choice set, any multiple choice elicited after would exhibit smaller chosen sets on average. We kept the results obtained with the 0 and 1-correspondence when it was clear that the priming did not happen.

We need choices from all possible subsets of the grand set of alternatives to be able to falsify the different models we will consider. We are restricted to small sets of alternatives, as the cardinal of the powerset grows exponentially with the size of the set of alternatives. For four alternatives, we have to study 11 choices, for five, 26 choices, and six, 57. In practice, X could contain 4 or 5 alternatives. We wanted the subjects to repeat the same choices for at least two different payments, and thus we settled on a set of 4 alternatives.

Finally, we wanted to study indifference in practice. We suspected that the amount of information

provided would influence the amount of indifference.⁷ Indeed, it is likely that if the information is scarce, it will be hard for subjects to establish the value of the alternatives, and the bonus payment may have a more substantial influence on their choices.

The experiment has been carried out in the Laboratoire d’Economie Expérimentale de Paris (LEEP), using zTree (Fischbacher (2007)). Subjects were recruited using Orsee (Greiner (2015)). All the sessions were in French, and subjects were paid in Euro. The show-up fee was 5€, and the average total gain was 10.28€. The experiment lasted between 40 and 60 minutes, depending on the different treatments and the speed of the subjects.

4.2.1 Tasks

Subjects chose between four different paid tasks (screenshots of the tasks are in Appendix C.1.1):

- An *addition* task, where subjects had to perform as many additions of three two-digit numbers as possible. They earned 30 cents for each correct sum.
- A *spell-check* task, where subjects faced a long text with spelling and grammar mistakes.⁸ They earned 10 cents for each mistake corrected and lost 10 cents for each mistake added. Their earnings were floored at 0 so that they could not lose money in this task.
- A *memory* task, where sequences of letters blinked on the screen and stopped after a random number of letters. Subjects had to give the three last letters that appeared on the screen. They earned 30 cents for each correct sequence.
- A *copy* task, where a large number of sequences of 5 letters appeared on the screen. Subjects had to copy the sequences. They earned 10 cents for each sequence.

Tasks involve some effort, which might influence the valuation made by the subjects. For instance, if they thought that the effort is more important in the memory task than in the spell-check task, and they expect the same gains, they might choose the latter over the former. It is not a problem *per se*, as long as it shapes their choices in the same way. Indeed, we chose effort in part because Augenblick, Niederle, and Sprenger (2015) claimed that real effort tasks induce sharper preferences, as subjects might feel strongly about some tasks. It is corroborated by their answers in a non-incentivized questionnaire administrated at the end of the experiment.

The effort might be a problem, in particular, in the high gain treatment. It may be seen in a first approximation as a discounting of the money increment depending on the effort involved, thus

⁷Another driver is the potential for different characteristics of the objects chosen to conflict (i.e., multidimensional choice), for instance, in the choice between different smartphones or different cars.

⁸For the interested French-speaking readers, it was the famous “dictée de Mérimée” with the modernized orthography of 1990: https://fr.wikipedia.org/wiki/Dictée_de_Mérimée The videos shown and the programmed tasks are available by asking the author.

lowering the increment of payment perceived for each new success in a task. It might induce some subjects to include a worse task in terms of monetary payments, especially in the high gains, if they perceive the discounted increment in gains as lower than the bonus payment of choosing an additional task. It is one reason why we do not include the high gain treatment in the principal analysis. We believe, however, that the low gains are so low that this effect should not influence them.

The whole choice process consisted of selecting tasks. At the end of the session, subjects had three minutes to earn as much as possible performing one task. We selected the task they performed by drawing one of the 33 choices they made at random. From this chosen set, we uniformly drew one task.

Additions and sequences were randomly generated and thus did not have an end. We told subjects that it was not possible to finish the spell-checking task in less than three minutes – and indeed, none did. Before performing the paid tasks, subjects always could train for at least 30 seconds, in order to get familiar with the interface. The training was, except for one information treatment, always done after they had made all their choices.

4.2.2 Timing

Subjects in the experiment went through five steps, which we explain in the next subsections. First, we read the instructions about the experiment to the subjects. One example is translated in Appendix C.1.2. The instructions are available (in French) here <http://www.bouacida.fr/files/eliciting-choice-correspondences/instructions/>. This part included a description of the tasks they had to choose. The descriptions differed across treatments. Each subject also had a printed version of the instructions in their cubicles. Second, subjects chose three times eleven choices, according to pay-for-certainty, at different payment levels. For some, we replaced the first payment by forced single choices. Third, we measured the subjects' risk-aversion, following Dohmen et al. (2011) method. Fourth, subjects answered a questionnaire on some socio-economic variable and their choices. Finally, subjects performed the task that had been selected and received their payments for the experiment afterward. The payment of subjects was made by drawing one of the 33 choices and one of the lotteries in the risk aversion elicitation. It is mostly in line with what Azrieli, Chambers, and Healy (2018) show to be incentive-compatible. Indeed, the payments for the lottery and the task are entirely independent and are likely to satisfy the no complementarities at the top hypothesis needed for incentive-compatibility.

4.2.3 Information Treatments

To investigate the influence of the information provided on the size of the chosen sets, we varied the explanations of the tasks. We have always given the explanations before the choices. Each subject faced one of the three possible treatments. In the *sentence* treatment, subjects received a vague description of the tasks, close to the description given in Section 4.2.1. In the *video* treatment, subjects first received the sentence treatment, and then watched a video explaining each task. The video showed the interface of the task and explained how to perform it.⁹ The forced single choice elicitation followed this treatment. Finally, in the *training* treatment, subjects first went through the video treatment. Then they trained on each task for 1 minute. This training happened *before* choosing. The quantity of information orders treatments: the sentence treatment is strictly less informative than the video treatment, itself strictly less informative than the training treatment.

4.2.4 Choices

We investigate choices with pay-for-certainty at three different payment levels and compare them to choices with a choice function. The difference between each set of 11 choices is the payment for adding an alternative. We studied three different bonus payment levels, no (0 cent), low (1 cent), and high (12 cents) gains.

Following pay-for-certainty, in each set, choosing an alternative implied a gain of $\frac{1}{|S|}\varepsilon$ where $|S|$ is the size of the choice set. For instance, if the choice set is of size two and the subjects faced a high gain, choosing an alternative pays 6 cents. Fractions of cent were paid by randomization: 0.25 cent corresponds to a 25% probability of getting 1 cent and a 75% probability of getting 0 cent. When the subjects chose several alternatives in a set, the computer used the uniform distribution to select the alternative eventually given to the subjects. The 11 choices at a given level of payment were performed in a row to avoid confusion between the different bonus payment levels. The order of the different bonus payment levels was random. The order of the different choice sets was random.

For each choice set, choosing meant saying “yes” or “no” to each task. Subjects had to choose at least one task. The order of alternatives shown on the screen was random. Once the subject chose, a confirmation screen appeared. It displayed the chosen alternatives and the associated gain. We did that to decrease the risk of errors from choosing hastily. Screenshots of a choice screen followed by a confirmation screen are in Figures 4.1 and 4.2. Subjects were reminded of the selection mechanism on each screen.

⁹The text used for spell-checking was different in the explanation and the real task.

Each selected task earns **6** cents.

If the computer draws this period at the end of the session, you will perform one of the tasks you are selecting. The computer will draw it at random among the tasks you are selecting. Each selected task is equally likely to be drawn.

You earn **6.00** cent(s) by selecting a task.

Select Addition? No
 Yes

Select Spellcheck? No
 Yes

Choice Made

Figure 4.1: A choice screen.

Each selected task earns **6** cents.

If the computer draws this period at the end of the session, you will perform one of the tasks you are selecting. The computer will draw it at random among the tasks you are selecting. Each selected task is equally likely to be drawn.

You might perform one of the following tasks

Addition	
Selecting these tasks will add to your payment (in cents)	6.00

Reminder: Only the gain of the period drawn at the end will be added to your final payment.

I confirm my choice | I wish to change my choice

Figure 4.2: A confirmation screen.

4.2.5 Questionnaire

Finally, subjects faced a non-incentivized questionnaire. The questionnaire investigated some socio-economic characteristics: gender, age, level and kind of education, and jobs. The answers are manually encoded using the National Institute of Statistics and Economic Studies of France classification, the Nomenclatures des Spécialités de Formation (NSF) for education, and the Classification of professions and socio-professional categories of 2003 (PCS 2003) for the kind and level of activities.

4.2.6 Data

The sessions took place between the 14th of November 2017 and the 29th of May 2018. The earliest started around 11.00am and the latest finished around 6.30pm. The time the choice was made during the day varied, but it should not matter too much as we mostly compare within subjects. There is at least some anecdotal evidence that the time of the session influenced choices, however.¹⁰

51 subjects participated in the forced single choice elicitation, during three separate sessions. These subjects follow the video information treatment. After the forced single choice elicitation, 17 chose according to pay-for-certainty with no and low gains.¹¹ Results obtained on these 17 subjects are in Appendix C.6. They allow us to compare more directly the forced single choice with pay-for-certainty.

172 subjects participated in the pay-for-certainty elicitation procedure with no, low, and high gains. Among these subjects, 102 followed the video treatment, 33 the sentence treatment, and 37 the training treatment. We drop from the analysis sample subjects who chose everything all the time, as the experimental design had no bite on their behavior. It represents the removal of 9 subjects, 4 in the sentence treatment, 4 in the video treatment, and 1 in the training treatment.

All subjects of the sample have done the measure of risk aversion, the questionnaire, and the tasks. The demographics of the sample shows that it is neither a representative sample of the population nor a typical student pool. In the principal analysis, we use strict revealed preferences, using Definition 1.5.

4.2.7 Demographics Characteristics

The average age of the sample is 35. The youngest subject is 18 and the oldest 76. The majority (56.40%) of the sample is 30 or younger. Figure 4.3 shows the age distribution of the sample. Figures

¹⁰One subject said that it was just after lunch and she was tired so that she chose an easier task.

¹¹All did it, but in two sessions, the results obtained showed that the participants did not understand the instructions. We dropped them.

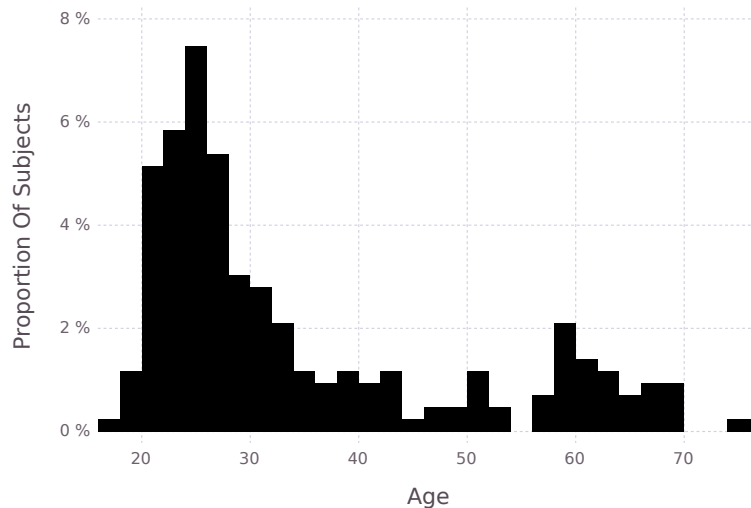


Figure 4.3: Age distribution of the sample.

4.4 and 4.5 show the different kinds of qualifications and studies. The majority of unemployed in the sample are students who do not work, as these two categories are lumped together by the French Statistical Institute.

The sample is almost gendered balanced. There are 53% of female in the sample. The level of education question only led to a response rate of 43%. Among those who answered, 51% have at least started a college education. Overall, the largest population of the sample is a student population.

4.3 Comparing Three Revelations Method

We first compare the possibility to choose multiple alternatives to the classical experimental method. We compare the results obtained in terms of consistency and preference revealed.

4.3.1 The Benchmark: Forced Single Choice

51 subjects were forced to choose a single alternative in all possible choice sets. Classical preferences rationalize 57% of observed choices. We assume that subjects reveal only strict preferences. As all revealed preferences are strict, each classical subject has six strict binary relations. We weaken this assumption in Appendix C.4 and show that more subjects are rationalized by classical preferences (80% to be precise) and that we can recover some indifference (29% of subjects have exactly one indifference relation, the others have only strict preference relations). We lose a lot of explanatory power, however, in allowing weak revealed preferences, as Appendix C.2 shows.

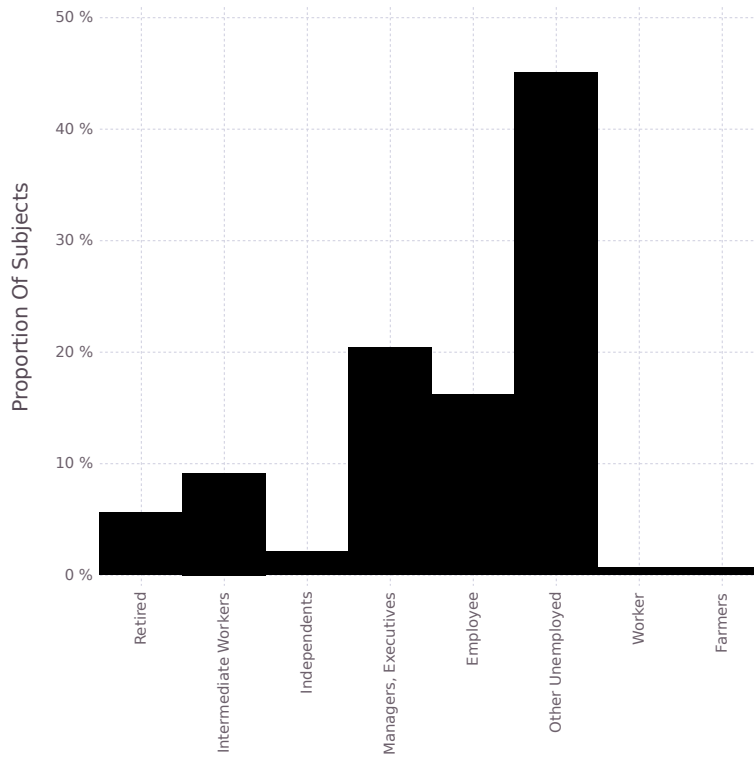


Figure 4.4: Professional occupation, when given (151 subjects).

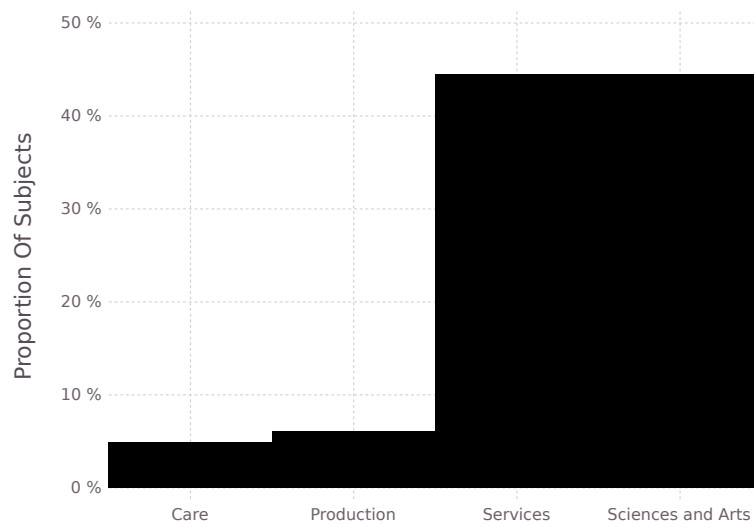


Figure 4.5: The domain of occupation, when given (172 subjects).

Table 4.1: Proportion of each task being maximal, restricted to subjects who are rationalized by classical preferences.

	Task				N ^a
	Addition	Spell-check	Memory	Copy	
Forced single choice	24%***	24%***	21%***	31%***	29
0-correspondence	51%*	51%*	52%	58%*	79
1-correspondence	72%	72%	65%	78%	72
Fully identified choice correspondence	66%	66%	53%	59%	32
Partially identified choice correspondence	50%*	50%*	59%	71%	66

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better. P-value of Fisher exact test with respect to the value given by the 1-correspondence. *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

^a In all the tables, N is the sample size.

A natural question from an individual welfare analysis is to ask what alternatives are preferred by decision makers. In this experiment, the fact that a subject prefers one task over another is of little relevance to the real world, but if we think about policy recommendation based on observed choices, this is a relevant question. Showing the distribution of the preference of decision makers does not yield a simple conclusion, as they are quite different. We can use a second best, however, which is to show the proportion of the sample with classical preferences that deem each alternative as one of the maximal ones. An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is strictly better. It allows us to build a crude order collective welfare order, by considering that overall, the best alternative collectively is the most approved. Table 4.1 shows precisely this. For choice functions, as we reveal only strict preferences, it means that each classical subject has exactly one maximal alternative, the one which is chosen in the grand set of alternatives. One limitation of this assessment is that we have to throw away all subjects who are not classical. Appendix C.7.2 provides a robustness check by using a non-preference based approach, using Condorcet winners. The results are mostly the same.

4.3.2 0-Correspondences

We study one of the simplest ways to allow subjects to choose several alternatives: pay-for-certainty with no bonus payment. The sample has 180 subjects. Classical preferences rationalize 44% of 0-correspondences. The proportion is lower than with choice functions, but the difference is not significant, with a Fisher exact p-value of 0.14. One benefit of allowing subjects to choose multiple alternatives is to observe indifference directly. The preference is only meaningful, so far, when subjects are rationalized by classical preferences. Figure 4.6 shows it is significant and heterogeneous. Some subjects are fully indifferent, whereas some have fully strict preferences. On average, 4.35

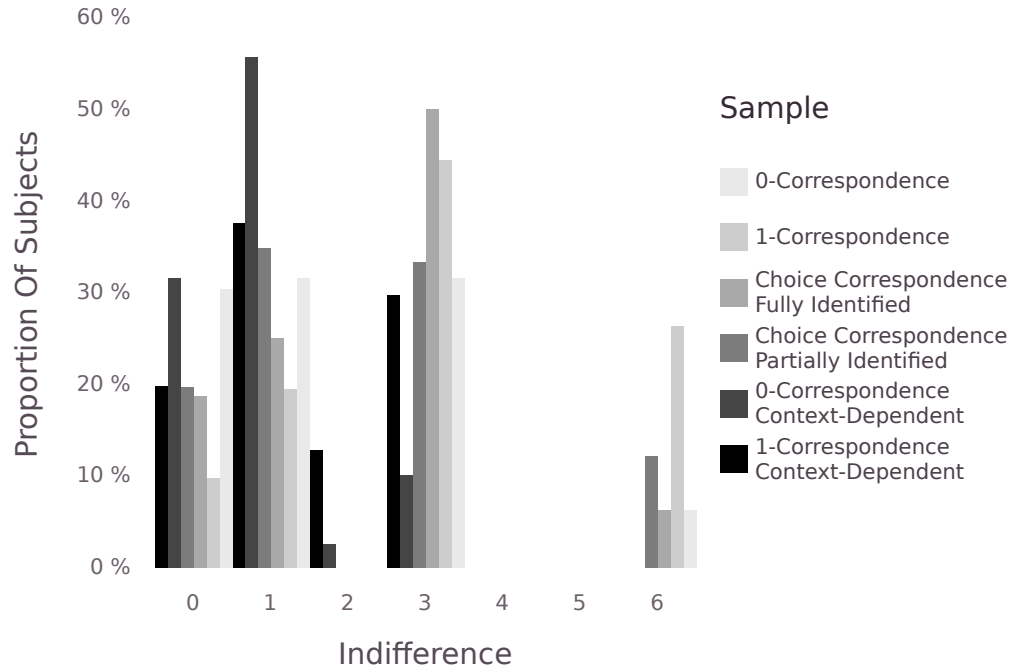


Figure 4.6: Histogram of the number of indifference relations with different samples.

relations are strict preference relations, and 1.65 are indifference relations.

Finally, Table 4.1 shows how much each alternative is maximal on average, for subjects who are rationalized by classical preferences. Because we have found significant indifference, some subjects might now have multiple maximal alternatives. It is what we observe, as we see that all alternatives are maximal for more than half of the sample, which is significantly higher than with forced single choice.¹² That is, for each alternative, a majority of subjects reveals it is maximal.

4.3.3 1-Correspondences

We interpret the 1-correspondences as the choice correspondence of the decision makers, thanks to the limit identification result of Proposition 3.2. We assume, therefore, in this subsection that the one cent bonus payment is low enough to warrant the limit interpretation. We will check this assumption in Section 4.3.4.

The sample has 180 subjects, which are the same as with 0-correspondences. Classical preferences rationalize 40% of 1-correspondences, which is significantly lower than with forced single choice.¹³ The difference between 0 and 1-correspondences is not significant, however.¹⁴ Figure 4.6 shows

¹²The p-value of the Fisher exact test between 0-correspondence and forced single choice are: addition: 0.022; spell-check: 0.022; memory: 0.006; copy: 0.021.

¹³The p-value Fisher exact test is 0.048.

¹⁴The p-value of the Fisher exact test is 0.52.

Table 4.2: Average number of strict preference relations and indifference relations for one subject with classical preferences.

	Strict Preference	Indifference	N
Forced Single Choice	6***	0***	29
0-Correspondence	4.35***	1.65***	79
1-Correspondence	3.11	2.89	72
Fully Identified Choice Correspondence	3.88**	2.12**	32
Partially Identified Choice Correspondence	3.92**	2.08**	66

Note:

P-value of two-sided two-sample t-test of equality with respect to the value given by the 1-correspondence. *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

that indifference is significant and heterogeneous, and higher than with 0-correspondences. The difference is significant, as the p-value of the Kolmogorov-Smirnov test of the two samples being drawn from the same distribution is lower than 0.001.¹⁵ Table 4.2 shows that on average, 3.11 relations are strict preference relations, and 2.89 are indifference relations. It is a significantly higher indifference than 0-correspondences

Finally, Table 4.1 shows how much each alternative is maximal on average, for subjects who are rationalized by classical preferences. As indifference is higher than previously, more subjects have multiple best alternatives, which we observe in aggregate, and the difference is often significant.

4.3.4 Identified Choice Correspondence

We can qualify the results obtained before by using the identification results based on set inclusion, rather than the limit. That is, we partially identify the choice correspondence of a subject when $c_0(S) \subseteq c_1(S)$ for all S , and we fully identify it when all the chosen sets are equal. In that case, we take into account two choices for each choice sets. The total sample is the same as before. We have to introduce some preliminary results, however, regarding the identification of the choice correspondence. Table 4.3 shows that we partially or fully identify the choice correspondence for a majority of subjects.

We can look at their distance from partial identification. Figure 4.7 represents the number of sets which violates $c_0(S) \subseteq c_1(S)$. All subjects whose choice correspondence is partially or fully identified never violates the above inclusion, which is why 58% of subjects are at 0. Most subjects whose choice correspondence cannot be identified are not far from identification. The median number of violating sets is 2. Some subjects, however, violate the set inclusion assumption radically. When the number of sets which violates $c_0(S) \subseteq c_1(S)$ is low, we might suspect that it is a mistake. In the

¹⁵All Kolmogorov-Smirnov test will be the same, so, from now on, we abbreviate it as the KS-test.

Table 4.3: Identification of choice correspondences.

	Identification		
	Full	Partial	None
Number of subjects	33	72	75
Proportion of subjects	18%	40%	42%

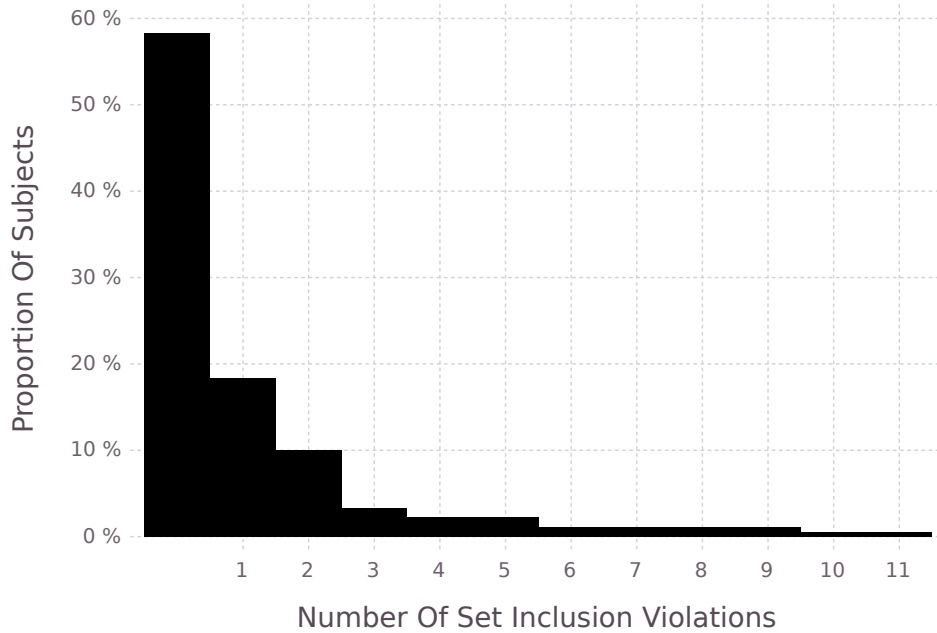


Figure 4.7: Histogram of the distance from partial identification. It shows the number of sets that violated set inclusion between the 0 and 1-correspondences.

remainder of this subsection, we exclude these subjects from the results and concentrate on fully and partially identified choice correspondences.

4.3.4.1 Fully Identified Choice Correspondences

In this subsection, we restrict the study to the sample of 33 subjects whose choice correspondence is fully identified. Classical preferences rationalize 97% of these subjects. It is significantly higher than with forced single choice and 0 and 1-correspondences, with p-values of Fisher exact test below 0.001.

We quantify the indifference and strict preferences of these subjects. Figure 4.6 shows a significant heterogeneity again in the kind of preference relations we observe. On average, 2.12 relations are indifference relations, and 3.88 are strict preferences, which is not significantly different from

0-correspondence.¹⁶ The distributions, however, are significantly different, with a p-value of KS-test lower than 0.001. The average number of indifference relations is significantly lower than 1-correspondences, and the distributions are also significantly different. The p-value of the KS-test is 0.04.

Finally, Table 4.1 shows how much each alternative is maximal on average for classical subjects. The figures are between 0 and 1-correspondences, which is expected, considering that the number of indifference relations is also in between.

4.3.4.2 Partially Identified Choice Correspondences

In this subsection, we restrict the study to the sample of 72 subjects whose choice correspondence is partially identified. We can use the partial identification results given in Section 3.3.3. We find that 92% of observed choices are compatible with classical preferences, i.e., may have a choice correspondence that satisfies WARP between their 0-correspondence and their 1-correspondence. This result is lower than with fully identified choice correspondences, but not significantly so, with a p-value of the Fisher exact test of 0.58. It is significantly higher than with forced single choice and 0 and 1-correspondences, however, with p-values of the Fisher exact test below 0.001.

When we use the classical compatible preference, remember that it is the one with the most strict preferences, as shown in Corollary 3.2, preferences relations exhibit both indifference and strict preferences. We see a significant heterogeneity in the preferences of subjects in Figure 4.6 again. On average, 2.08 relations are indifference relations, and 3.92 are strict preference relations, which is not significantly different from identified choice correspondences. The distribution of indifference is significantly different; the p-value of the KS-test is lower than 0.001.

Finally, Table 4.1 shows how much each alternative is maximal on average for classical subjects. The figures are significantly higher than with choice functions and not significantly different from 0 or 1-correspondences or fully identified choices correspondences.¹⁷

4.3.5 Related Literature and Discussion

Few papers proceed to similar tests of rationalizability by classical preferences on experimental data. The setups can be widely different, but their results are not. They all show significant violations of WARP. Except for Costa-Gomes, Cueva, and Gerasimou (2016), however, they all use choice functions to elicit preferences.

¹⁶The p-value of the two-sample two-sided unequal variance t-test of equality of the mean is 0.16.

¹⁷Using a Fisher exact test of equality for each task. For choice functions, the p-values are 0.03 for addition and spell-check and below 0.001 for memory and copy.

In Section 2.3.1, 47% of subjects satisfy WARP all the time in the experimental data of Manzini and Mariotti (2010), which is similar to the figure we found on choice functions. Choi, Fisman, et al. (2007) found around 35% of subjects violating WARP for choices over risky assets. In the closely related large-scale field experiment of Choi, Kariv, et al. (2014), around 90% of subjects violate WARP for a similar choice task. The subject pool here is in between their two pools, as shown in Section 4.2.7. Costa-Gomes, Cueva, and Gerasimou (2016) have two main treatments in their experiments: one where subjects are forced to choose one alternative and one where they can postpone at a cost. When forced to choose a single alternative, 54% of subjects satisfy WARP, and when not forced, 73% satisfy WARP. The first figure is remarkably close to ours, whereas the second is quite higher.

Comparing the results obtained, assuming the limit identification and the partial and full identification yield several conclusions. Limit identification has the benefit of assessing the consistency of the whole sample. It yields, however, a lower (potential) rationalizability with classical preferences, and a higher indifference, as shown in Table 4.2. If we believe that the sample selection with full and partial identification is not too severe, these results suggest that fractions of 1 cent were not low enough to use a limit identification result. The rest of the literature inclines us to think that the results obtained with full and partial identification are overly optimistic, however. We believe that the results obtained with 1-correspondences are more representative of what we should expect in terms of consistency in an experiment.

To the best of our knowledge, we are the first to investigate indifference directly in an incentive-compatible manner. We find that indifference is significant in both 0 and 1-correspondences. It might have consequences for collective welfare analysis, as aggregating preferences with some indifference is more accessible than when preferences are fully strict, as shown by Table 4.1. Assuming 1 cent is a low enough incentive not to bundle together strict preferences and indifference. The significant difference in the number of indifference relations between 0- and 1-correspondences cannot be reconciled with the idea that the whole population has a preference for randomization. Indeed, otherwise, all subjects who value randomization would already choose all the alternatives they are indifference in between with the 0-correspondence, which is not what we observe.

4.4 Going Further with Correspondences¹⁸

In addition to qualifying the robustness of the classical experimental method and the consistency of observed choices, our experiment aimed at testing models of decision making involving intransitive indifference and menu-dependent choices, as described in Section 1.2. These models rationalize the same observed choices as classical preferences do on choice functions. They are therefore indistin-

¹⁸We use correspondence to lump together 0-correspondences, 1-correspondences, partially and fully identified choice correspondence.

Table 4.4: Rationalizability by a weakening of classical preferences.

	Correspondence		Identified	
	0	1	Partially	Fully
Classical Preferences	44%	40%	92%	97%
Semi-Order	45%	42%	94%	97%
Interval Order	45%	42%	94%	97%
Partial Order	45%	42%	94%	97%
Occasional Optimality	57%*	57%**	97%	97%
Menu-Dependent	82%***	91%***	100%*	100%
Context-Dependent	88%***	96%***	100%*	100%
N	180	180	72	33

Note:

Significance levels are assessed by using a Fisher exact test.

The baseline value are figures obtained with classical preferences. *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

guishable on forced single choice and require at least a 0-correspondence elicitation to be explored. We start with menu-independent choice models and then explore menu-dependent choice models.

4.4.1 Intransitive Indifference

In theory, it is possible to rationalize choice correspondences when classical preferences fail with relaxations of complete and transitive preferences. We relax transitivity of the indifference using the axioms given in Section 1.2.1. We can investigate these models because we have observed menu choices and not only a singleton choice. To the best of our knowledge, we are the first to study the empirical validity of intransitive indifference in a systematic way.

We explore three intransitive indifference models: semi-order, interval order, and partial orders. The first two impose some consistency on the indifference and strict preference relations, whereas the latter only requires the strict preference to be transitive.

Table 4.4 summarizes the satisfaction of the different models depending on the specification of the choice correspondence considered. Using menu-independent weakening of classical preferences is not very helpful to rationalize correspondences in the experiment. The various weakening rationalizes marginally more correspondences than classical preferences, and the difference is never significant. Interestingly, partial orders, interval orders, and semi-orders rationalize the same correspondences in the experiment. It does not have to be the case in theory, as the difference in explanatory power in Appendix C.2 shows.

4.4.2 Menu-Dependent Choices

In addition to intransitive indifference, we can explore models of menu-dependent choice. Specifically, we explore the occasional optimality model of Frick (2016), and the menu- and context-dependent models of Aleskerov, Bouyssou, and Monjardet (2007). Table 4.4 show the menu- and context-dependent choices rationalize almost all 0- and 1-correspondences. It is significantly more than classical preference. When we consider partially or fully identified choice correspondences, the baseline obtained with classical preferences is much higher, and the results lose much of their significance.

One benefit of the context-dependent choice model is to show that one underlying preference can explain choices that are menu-dependent. It is quite easy to build this preference with the data of our experiment, as explained in Section 1.2.1.6. We consider the preferences obtained when we can rationalize choices with the context-dependent choice model only. The preferences revealed by the other models are the same than the context-dependent choice model but applied on a smaller sample. On average, subjects who are rationalized by the context-dependent choice model but not rationalized by classical preferences have 0.91 indifference relations on 0-correspondences and 1.52 on 1-correspondences, which is significantly lower than the similar figures obtained with classical preferences.¹⁹ Figure 4.6 show that the indifference revealed by the context-dependent choice is significantly lower than with classical preferences. The distributions are also significantly different, as the p-values of the KS-tests are below 0.001. Notice that subjects whose choices can be rationalized by classical preferences are excluded from the figures of the context-dependent choice, even though when a subject is classical, his choices can be rationalized by a context-dependent model. We did this in order to avoid double counting subjects who are rationalized by classical preferences.

Finally, the lower number of indifference relations of subjects which satisfy the context-dependent choice but not the classical model has a corollary. Fewer alternatives are maximal, as shown by comparing Tables 4.5 and 4.1. The differences with classical preferences on 0-correspondences are mostly insignificant, but they are significant on 1-correspondences.

4.4.3 Discussion

Using correspondences, we have explored revealed preferences that are not classical. We have shown in Section 4.3 that allowing subjects to choose several alternatives decreased their rationalizability by classical preferences. Intransitive indifference and menu-independent models do yield to significantly higher rationalizability in this experiment. On the flip-side, however, we show in this section that we rationalize almost all correspondences with menu-dependent models. The difference in satisfaction

¹⁹The p-values of the two-sided two-sample unequal variance t-test of equality of the means are below 0.001 in both cases.

Table 4.5: Proportion of each task being maximal, restricted to subjects who are rationalized by context-dependent choice and not classical preferences. In parenthesis are the p-value of the Fisher exact test of equality with classical preferences, task by task and ε -correspondence by ε -correspondence

	Task				N
	Addition	Spellcheck	Memory	Copy	
0-Correspondence	44% (0.524)	35% (0.077)	37% (0.078)	48% (0.264)	79
1-Correspondence	51% (0.009)	39% (<0.001)	50% (0.075)	58% (0.012)	101

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better. No difference are significant, according to a Fisher exact test.

between WARP on forced single choice and FP on 0 and 1-correspondences is significant.²⁰ In that sense, we have restored the rationalizability of subjects.

Additionally, the context-dependent models allow us to identify a unique underlying classical preference explaining the choice. We study the preferences obtained when the context-dependent model rationalizes the observed choices of the subjects. They exhibit significantly less indifference, which implies that the samples of subjects satisfying WARP and the sample of subjects which satisfy FP but not WARP are different, not only on their rationalizability but also on their preferences.

4.5 Influence of the Information Provided

In this experiment, we varied the information provided to the subjects. We expect the information provided to influence the choice of decisions makers, as it is likely to influence their valuation of each task. 115 subjects took part in the video treatment, 29 in the sentence treatment, and 36 in the training treatment. The sample sizes are small for the latter two, rendering any exploitation of full and partial identification results hardly meaningful so that we only consider 0 and 1-correspondences.²¹ We will not show results on these.

We assess how the information provided influences the consistency of subjects. Table 4.6 summarizes the satisfaction of the different models depending on the information provided, on 0- and 1-correspondences.²² Overall, it looks as if the information provided is slightly harmful in terms of consistency, in particular on 1-correspondences. The differences in the different axioms satisfaction between information treatments are not significant, according to Fisher exact tests.

²⁰The p-values of the Fisher exact test are both below 0.001.

²¹In the sentence treatment, we fully identify the choice correspondence of 4 subjects and partially identify it for 19. In the training treatment, figures are respectively 9 and 12.

²²Results for semi-order and interval order have been removed from the analysis, as they provide no additional

Table 4.6: Rationalizability of 0 and 1-correspondences, depending on the information provided.

	0-Correspondence			1-Correspondence		
	Sentence	Video	Training	Sentence	Video	Training
Classical Preferences	55%	40%	47%	52%	39%	33%
Partial Order	59%	41%	47%	52%	42%	33%
Occasional Optimality	66%	57%	53%	69%	52%	61%
Menu-Dependent	83%	79%	92%	100%	88%	94%
Context-Dependent	93%	85%	92%	100%	94%	100%
N	29	115	36	29	115	36

Table 4.7: Average gains in the tasks, depending on the information provided and the model rationalizing choices. Context-Dependent choices does not include classical preferences.

	Information		
	Sentence	Video	Training
Classical Preferences	2.63€*	3.08€	3.85€
Context-Dependent	2.52€**	2.84€**	3.78€

Note:

Significance levels are reported with respect to the training treatment, for a two-sided two-sample unequal variance t-test of equality of the means. *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

On the other hand, Table 4.7 shows that the information provided is beneficial in terms of gains in the tasks. The difference between the training and the two other treatments are significant, but the difference between the sentence and the video treatment are not, maybe because of the small sample sizes.

Turning ourselves to the analysis of the preferences revealed by the subjects, we see that indifference is heterogeneous between subjects, and the information treatments indeed influence the amount of indifference. Figure 4.8 shows that some subjects are entirely indifferent, whereas some have only strict preference relations. Overall, more information leads to less indifference.

We investigate the cross-influence of the information treatment and the bonus payment levels jointly by running an ordinary least-square regression with and without subjects fixed effects. We use interaction effects between the level of bonus payment and the information treatment to capture the idea that the influence of the bonus payment may depend on the information provided.

information compared to partial orders.

Table 4.8: OLS regressions on the number of indifference relation on the sample of subjects rationalizable by context-dependent choice, controlling for rationalizability by classical preferences.

	Indifference	
	(1)	(2)
Intercept	0.63** (0.215)	
Bonus	1.93*** (0.367)	1.76*** (0.363)
Video	0.11 (0.248)	
Training	0.09 (0.328)	
Video & Bonus	-0.99* (0.419)	-0.97* (0.393)
Training & Bonus	-1.47** (0.496)	-1.31** (0.406)
Classical	1.13*** (0.169)	1.13*** (0.206)
Fixed Effects	No	Yes
N	331	308
Adjusted R ²	0.24	0.79

Note:

Fixed effects are at the subject level. Robust standard deviations are in parenthesis. *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

$$I = \beta_1 \text{bonus} + \beta_2 \text{training} + \beta_3 \text{video} + \beta_4 \text{training} \times \text{bonus} + \beta_5 \text{video} \times \text{bonus} + \beta_6 \text{WARP} + \varepsilon \quad (4.1)$$

Table 4.8 shows the coefficient of the regressions given in Equation (4.1). *Bonus* is a dummy of value one for 1-correspondences and zero for 0-correspondences. *Training* is a dummy of value one if it is the training treatment. *Video* is a dummy of value one if it is the video treatment. *WARP* is a dummy of value one when the decision maker is rationalized with classical preferences, and of value zero otherwise. The baseline treatment considered is, therefore, the sentence treatment with no bonus payment as the information is fixed at the subject levels, no direct effect of information that can be captured in the fixed effect regression.

The regression with fixed effects shows that there is a positive and significant impact of the gain

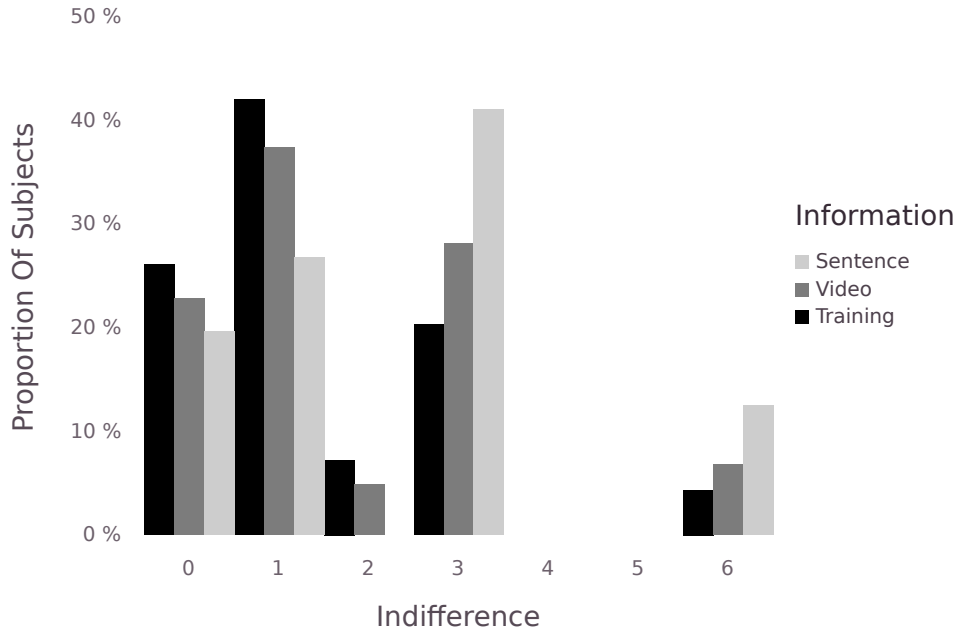


Figure 4.8: Histogram of the number of indifference relations by information provided. The p-value of the KS-test is equal to 0.01 between the training and the sentence treatment, and below 0.001 between the video treatment and the other two.

level on the number of indifference relations, as going from 0 to 1 increases on average the number of indifference relations by 1.75. The cross effects of information and the bonus payments are negative and significant and partially cancels out the direct effect of the bonus payments. The more information is provided at a given bonus payment and the lower the number of indifference relation is. The fact that a decision maker can be rationalized with classical preferences also has a positive and significant effect on the number of indifference relations. It indicates that classical and non-classical decision makers are likely to be different. The adjusted R^2 is much higher with fixed effects, indicating that much of the explanatory power is due to subjects heterogeneity. Overall, the small bonus of 1 cent is enough to drive the choice of the subjects, when they lack information about the tasks.

Finally, we can look at which alternatives the subjects deems as the best one, depending on the information provided. We observe a clear downward trend in the proportion of subjects who have memory as a maximal alternative, as the quantity of information provided increases. The other alternatives do not vary significantly with the information provided. A regression analysis (in Appendix C.7.1) confirms that only the choice of the memory task is influenced on aggregate by the information provided. It is in line with the results given before on how indifference varied with the information provided.

Table 4.9: Proportion of each alternative being maximal, depending on the information provided. The sample is made of all 0 and 1-correspondences which can be rationalized with a context-dependent preference. The results do not change much depending if we restrict to classical preferences.

	Task				N
	Addition	Spellcheck	Memory	Copy	
Sentence	59%	50%	75%**	59%	56
Video	55%	45%	52%	58%	206
Training	46%	57%	26%***	67%	69

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better. Significance levels are given with respect to the video treatment, column-wise, using a Fisher exact test.

*: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

4.6 Conclusion

In this chapter, we have implemented pay-for-certainty with a uniform selection mechanism in a laboratory experiment. The method is working in practice, as a majority of subjects are consistent with partial or full identification with set inclusion, the main requirement for its validity. It is also valuable, as it confirms the idea that restricting the choice of decision makers to one alternative in practice is a real constraint, at least for some applications.

We have shown that when we partially or fully identify the choice correspondence of decision makers, we cannot falsify the fact that they have classical preferences. If we use the limit identification result, on the other hand, classical preferences rationalize less than with forced single choice. The benefit of getting sets that are chosen rather than a single alternative is that we can use the context-dependent model to rationalize most of the 1-correspondences. The difference between classical preferences and context-dependent choice imply that many subjects behave as an approximate maximizer. They do not seem to attribute a precise enough value to each task to maximize their preferences in a classical sense. They behave consistently enough for us to reveal a preference with correspondences, however.

The preferences revealed in this experiment with choice correspondence exhibit widespread indifference. It is significantly higher than what could have been obtained with choice functions, even assuming weak revealed preferences. It does not disappear even when there are no incentives to choose several alternatives, nor when the information provided is quite complete. Even a simple elicitation method for choice correspondence, i.e., 0-correspondence elicitation, which does not identify the choice correspondence alone, is better than choice functions to study indifference. It has important implications for individual welfare analysis, as it implies that we overestimate strict pref-

erences, and thus overestimate the individual loss due to switching to a seemingly Pareto-dominated alternative. It also has collective welfare implication, as the preferences with some indifference may break Arrow's impossibility theorem.

Chapter 5

Conclusion

This dissertation takes a look at the link between choices, preferences, and individual welfare. We aimed at extending the classical revealed preference paradigm, by going beyond rationalizability with classical preferences, while keeping a relatively traditional revealed preference understanding of choices. In Chapter 2, with Daniel Martin, we have shown the value of new approaches from behavioral welfare economics in practice. They allow us to make relatively precise welfare recommendation.

In Chapter 3, we have built and characterized a method that bridges the gap between the empirical and the theoretical literature on revealed preferences. We characterize a method to identify choice correspondences in practice. It allows us to adopt a less restrictive understanding of rationality in practice, in Chapter 4. We show that most subjects can be understood as approximate utility maximizers, and in this sense we restore their rationality.

The results shown in this dissertation leaves some open questions. A question mark regarding welfare analysis is the validity of the revealed preferences in Chapters 2 and Chapter 4. We do not know, and it is hard to know if the revealed preferences obtained reveal the underlying welfare order of the decision makers.

Additionally, each chapter leaves some open questions. The external validity of the results obtained in Chapter 4 is not very clear. Tasks were built for this experiment, and an experiment with more common alternatives would be insightful and could confirm or infirm the results of the experiment. One weakness of the method presented in Chapter 3 is its incompatibility with lotteries. It means that one of the most widely used choice objects in experiments cannot be used with pay-for-certainty.

Appendix A

Appendix of Chapter 2

A.1 Choice Set Size

Figure A.1 shows that the choice set sizes in the consumption data appear to be close to uniform. However, a KS-test of the sample being drawn from a uniform distribution rejects this possibility with a p-value lower than 0.001.

A.2 Correlations for Uniform Random Demands

In order to check the robustness of the results found in Section 2.3.4, we ran a robustness check using uniform random demands.

A.2.1 Uniform Random Methodology: Consumption Data

Our primitives are quantities (for each individual and each period) and prices (for each market and each period). We first compute the observed expenditures for each period. We then build vectors of quantities that constitute our random bundles. Each bundle is built by drawing a vector from the simplex uniformly¹ and then multiplying it by the observed expenditures of each period to get (random) quantities. We now have observed prices and random quantities for each individual. We simulate each observed individual 1,000 times. In total, we have 1,190,000 simulated individuals.

¹Using a flat Dirichlet distribution (see https://en.wikipedia.org/wiki/Dirichlet_distribution)

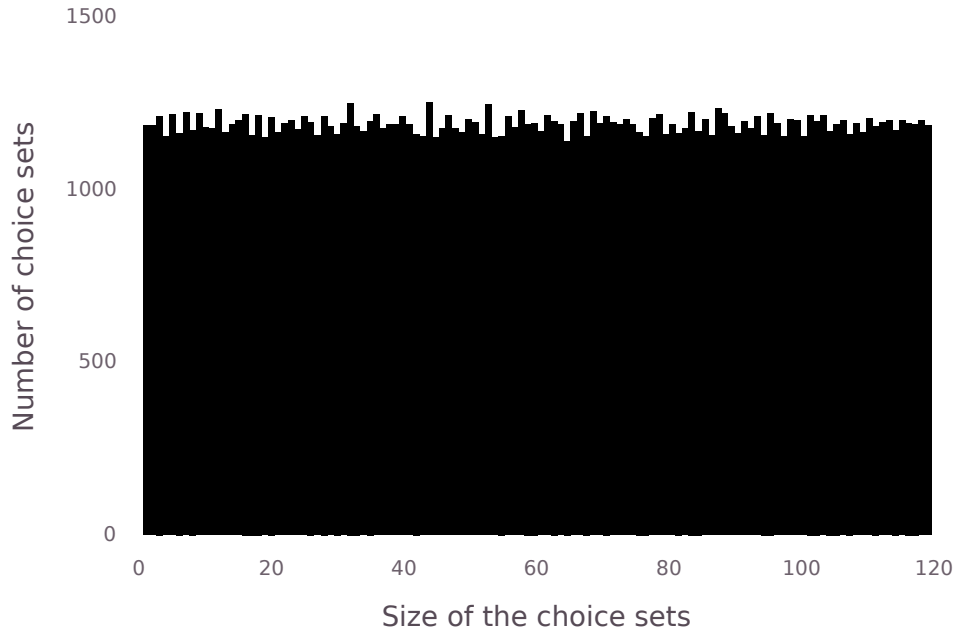


Figure A.1: Histogram of the size of choice sets in the consumption data.

A.2.2 Uniform Random Methodology: Experimental Data

Our primitives are the twelve possible choice sets. In each choice set, each alternative has an equal probability of being chosen. Here, there is no difference between individuals, as they all faced the same choice sets. In total, we have 1,000,000 simulated individuals.

A.2.3 Results

We find that uniform random demands are much less consistent than observed demands. In the experimental data, all but one simulation has RP cycles. In the consumption data, 100% of the simulations have RP cycles, 75% have SUCR cycles, and 0% have TC cycles. We restrict the sample to simulations with RP cycles. For these simulations, SUCR and TC also less complete. In the experimental data, 93.85% of simulations are a half or less of classical preference with SUCR and 97.65% with TC. In the consumption data, on average, SUCR is 97.86% of a classical preference, whereas TC is 94.98%, which is much lower than in the observed data.

The average predictive power drops as well. In the experimental data, the average value of Selten's index is 0.22 for SUCR and 0.16 for TC. In the consumption data, the average Selten's index is 0.98 for SUCR and 0.96 for TC. The correlation with our proposed measures are presented in Table A.1, and the figures are qualitatively similar to those of Table 2.8. It suggests that the relationship between the number and the fraction of direct RP cycles and the predictive power of SUCR and TC holds more generally than just the demands we observe in our data sets.

Table A.1: Summary of correlations with the average Selten's index at the simulation level (for simulations with RP cycles).

	Selten's Index					
	Experimental		Consumption		Acyclic SUCR	
	SUCR	TC	SUCR	TC	SUCR	TC
Number of relation elements	0.96	0.97	0.58	0.88	0.56	0.83
Number of direct RP cycles	-0.96	-0.58	-0.64	-0.84	-0.63	-0.78
Number of length 3 RP cycles	-0.89	-0.74	-0.49	-0.75	-0.49	-0.68
Number of length 2, 3, and 4 RP cycles	-0.91	-0.69	-0.38	-0.61	-0.40	-0.57
Directness index (DI)	0.54	0.81	0.47	0.70	0.41	0.61

Appendix B

Appendix of Chapter 3

B.1 Own Randomization

We have run a small-scale experiment using the own randomization selection mechanism. 37 subjects participated in this experiment. Most of the experimental design is similar to what was presented in Section 4.2. In particular, the tasks were the same and were described following the video treatment. In the first step of the experiment, shown in Figure B.1, they were asked to choose a probability distribution over the alternatives they faced, in the 11 possible choice sets. In a second step, shown in Figure B.2, we elicited their willingness to pay 10 cents to get the distribution they had chosen over the uniform distribution, in the 11 sets.

One of the 22 choices made was implemented to assign them a task. If it was one of the first 11 choices, the task was drawn using their distribution. If it was one of the last 11 choices, first we look at what distribution they had chosen between the uniform and their own and then assign the task using the chosen distribution. If they chose the uniform, they earned an additional 10 cents. We framed it at a cost and gave them a show-up fee of 5.10€ to be in line with the experiment of Chapter 4. It had to be a cost to elicit the indifference properly here.

We have computed their choice correspondence assuming that any alternative with positive probability is a chosen alternative. Efe A Ok and Tserenjigmid (2019) introduced two notions linking rationality of stochastic choice and deterministic choice. Here, we are using what they called *upper rationality*. The sizes of the chosen sets are in Table B.1. Only 8% of the subjects choose Bayesian distributions, using the distribution over the four alternatives as a baseline. It is clear here that chosen set sizes are on average much larger than with the uniform selection mechanism, casting doubt about the validity of this experimental design or the own randomization selection mechanism.

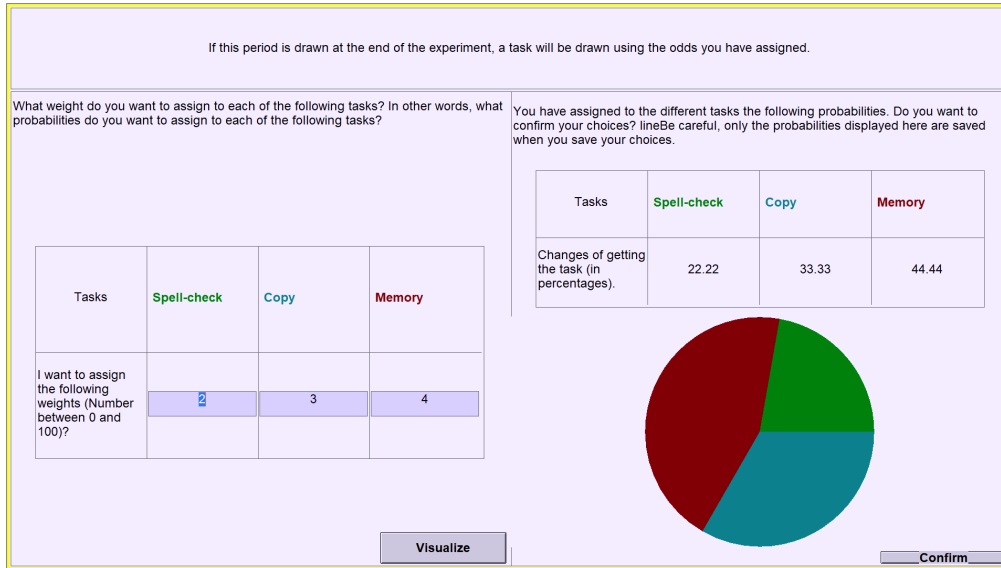


Figure B.1: Choosing a distribution.

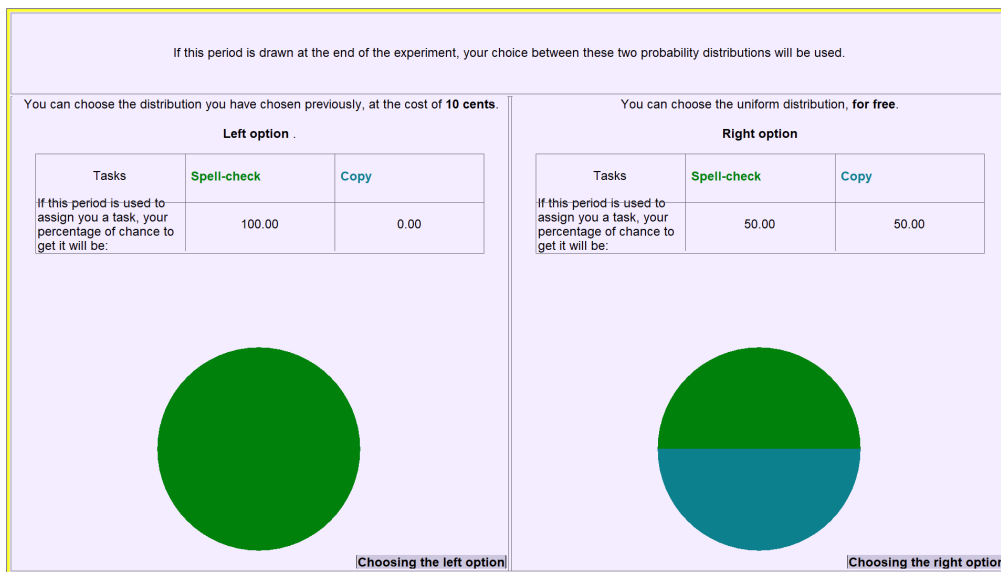


Figure B.2: Choosing between a chosen distribution (left) and the uniform distribution (right).

Table B.1: Size of the chosen set when subjects chose their distribution over the alternatives.

Set size	1	2	3	4
2	21%	79%		
3	13%	20%	67%	
4	5%	8%	22%	65%

B.2 Extending Pay-For-Certainty to Infinite Sets

We discuss here why extending pay-for-certainty to infinite sets is not straightforward. In particular, why the Andreoni and Sprenger (2012) method for preferences over risky assets cannot be used with pay-for-certainty.

B.2.1 Infinite Choice Sets

Take (X, μ) to be a measurable set. The natural extension of pay-for-certainty in this context is:

- The decision maker can choose any subset of $S \subseteq X$;
- The gain of choosing one subset is $\left(1 - \frac{\mu(c(S))}{\mu(S)}\right) \varepsilon$;
- The alternative given to the decision maker is selected from the chosen set using the uniform selection mechanism.

This procedure is the natural counterpart of the pay-for-certainty with the uniform selection mechanism on finite sets. It reduces to pay-for-certainty on finite sets using the uniform measure.

This naive counterpart does not work as intended, as an example illustrates.

Example: *Infinite choice set* Take $X = [0, 1]$, with the canonical measure on \mathbb{R} . The decision maker has to choose the certainty equivalent in $[0, 1]$ of the lottery $(1/2, 0; 1/2, 1)$. Assume that she is imprecise about her exact certainty equivalent but knows it is in the interval $I = [0.4, 0.5]$. The gain from this choice is:

$$\left(1 - \frac{\mu([0.4, 0.5])}{\mu([0, 1])}\right) \varepsilon = \left(1 - \frac{.1}{1}\right) \varepsilon = 0.9\varepsilon$$

By density of \mathbb{Q} in \mathbb{R} , an equivalent choice is to choose the set $I' = \mathbb{Q} \cap I$, a set of measure 0, which yield to a gain of ε .

Avoiding this pitfall requires forcing the choice of convex sets in S , which in turns requires S to be a convex set. It is an additional technical restriction on the choice of the decision maker that has no clear justifications.

B.2.2 Finite Choice Sets on Infinite Sets

One might want to restrict to finite choice sets with an underlying infinite structure. It does not work straightforwardly either. Following the previous example, but now restricting the choice of the certainty equivalent to be one or several of the following values $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.

An interpretation problem arises: what happens if his certainty equivalent is 0.41? Assuming that ε is small enough, would he choose 0.4 or $\{0.4, 0.5\}$? It is not clear. If she chooses 0.4, what is his certainty equivalent? We certainly cannot guarantee that it is equal to 0.4; it may as well be 0.42 or 0.38.

The pay for certainty procedure can be applied on finite sets with no underlying infinite structure (i.e., not for risky lotteries, not for consumption choices). The extension to a choice set with an underlying infinite structure is not straightforward and requires additional assumptions.

Appendix C

Appendix of Chapter 4

C.1 Design

C.1.1 Tasks

Subjects chose between four different paid tasks. The programs to run the tasks and the videos used in the video and training treatments are available from the author, on request. The tasks were:

- An *addition* task, where subjects had to perform as many additions of three two-digit numbers as possible. They earned 30 cents for each correct sum. A screenshot of one addition is available in Figure C.1.
- A *spell-check* task, where subjects faced a long text with spelling and grammar mistakes. They earned 10 cents for each mistake corrected and lost 10 cents for each mistake added. The text they faced (in French), as well as the interface of the task, is in Figure C.2.
- A *memory* task, where sequences of blinking letters appeared on the screen and stopped after a random number of letters. Subjects had to give the three last letters that appeared on the screen. They earned 30 cents for each correct sequence. The interface to give the sequence of blinking letters is shown in Figure C.3.
- A *copy* task, where a large number of sequences of 5 letters appeared on the screen. Subjects had to copy the sequences. They earned 10 cents for each sequence. The interface of the task is in Figure C.4

At the end of the tasks, subjects had feedback on their gains. In the training part, it was on their potential gains. In the training treatment, when subjects performed the tasks before choosing, they also received feedback on their performance.

Tâche

Nombre de bonnes réponses : 0

Question N°1 : **93+52+97 =** Question suivante >

Il vous reste 1 minute(s) et 17 secondes

Figure C.1: Addition task.

Texte initiale à corriger : Score : 63

Pour ~~parler~~ parler ~~een~~sans ~~ambiguité~~ambiguïté, ce dîner ~~è~~ à Sainte-Adresse, ~~près~~ du Havre, malgré les ~~effluves~~effluves ~~embaumés~~embaumés de la ~~mer~~mer, malgré les vins de très ~~bons~~ bons ~~et~~ crus, les ~~cuisseaux~~cuisseaux de veau et les ~~cuisseaux~~cuisseaux de chevreaux ~~prodigés~~prodigués par l'~~amphitryon~~amphitryon, fut un vrai guépié.

Quelques ~~soient~~soient, et quelques ~~exigus~~exigus ~~qu'ait~~qu'ait pu paraître, à côté de la somme ~~dûe~~due, les ~~arrhes~~arrhes ~~qu'étaient~~qu'étaient censés avoir donné la ~~douairière~~douairière et le marguillier, il était ~~infâme~~infâme d'en vouloir pour cela à ces fusilliers jumeaux et mal bâtis, et de leur ~~infliger~~infliger une ~~raclée~~raclée, alors ~~qu'ils~~qu'ils ne ~~songeaient~~songeaient qu'à prendre des rafraîchissements avec leur coreligionnière.

Quoi qu'il en soi, ~~ces~~ bien a tord que la douairière, par un contre sens exhorbitant, c'est laisser entraîner à prendre un rateau et qu'elle s'est cru obligé de frappé l'exigent marguillier sur son omoplate vieillie.

Deux alveoles furent brisées ; une disenterie se déclara suivie d'une fisie, et l'imbécillité du malheureux s'accrut.

- Par saint Martin ! Quelle hémorragie ! s'écria ce bêtire.

A cet événement, saisissant son goupillon, ridicule excédent de bagage, il la poursuivi dans l'église toute entière.

Veillez corriger le texte ci-dessous :

Pour parler sans ambiguïté, ce dîner à Sainte-Adresse, près du Havre, malgré les effluves embaumés de la mer, malgré les vins de très bons crus, les cuisseaux de veau et les cuisseaux de chevreaux prodigués par l'amphitryon, fut un vrai guépié.

Quelques soient, et quelques exigus qu'ait pu paraître, à côté de la somme due, les arrhes qu'étaient censés avoir donné la douairière et le marguillier, il était infâme d'en vouloir pour cela à ces fusilliers jumeaux et mal bâtis, et de leur infliger une raclée, alors qu'ils ne songeaient qu'à prendre des rafraîchissements avec leur coreligionnière.

Quoi qu'il en soi, ces bien a tord que la douairière, par un contre sens exhorbitant, c'est laisser entraîner à prendre un rateau et qu'elle s'est cru obligé de frappé l'exigent marguillier sur son omoplate vieillie.

Deux alveoles furent brisées ; une disenterie se déclara suivie d'une fisie, et l'imbécillité du malheureux s'accrut.

- Par saint Martin ! Quelle hémorragie ! s'écria ce bêtire.

A cet événement, saisissant son goupillon, ridicule excédent de bagage, il la poursuivi dans l'église toute entière.

Il vous reste 0 minute(s) et 6 secondes

Figure C.2: Spell-check task.

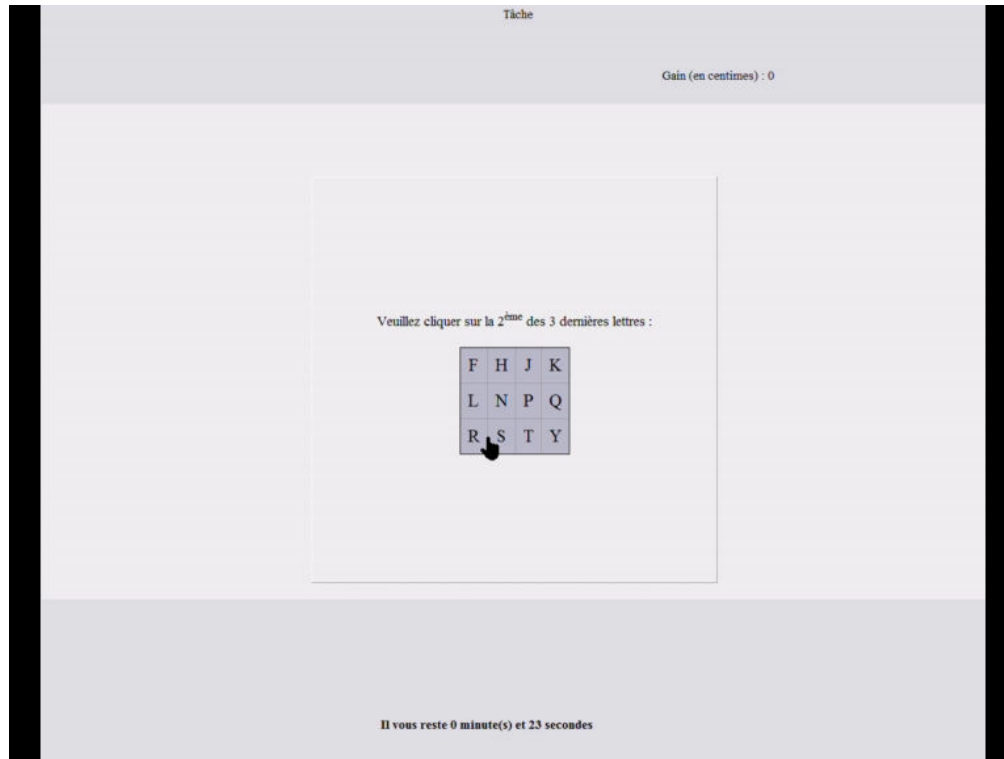


Figure C.3: Memory task.

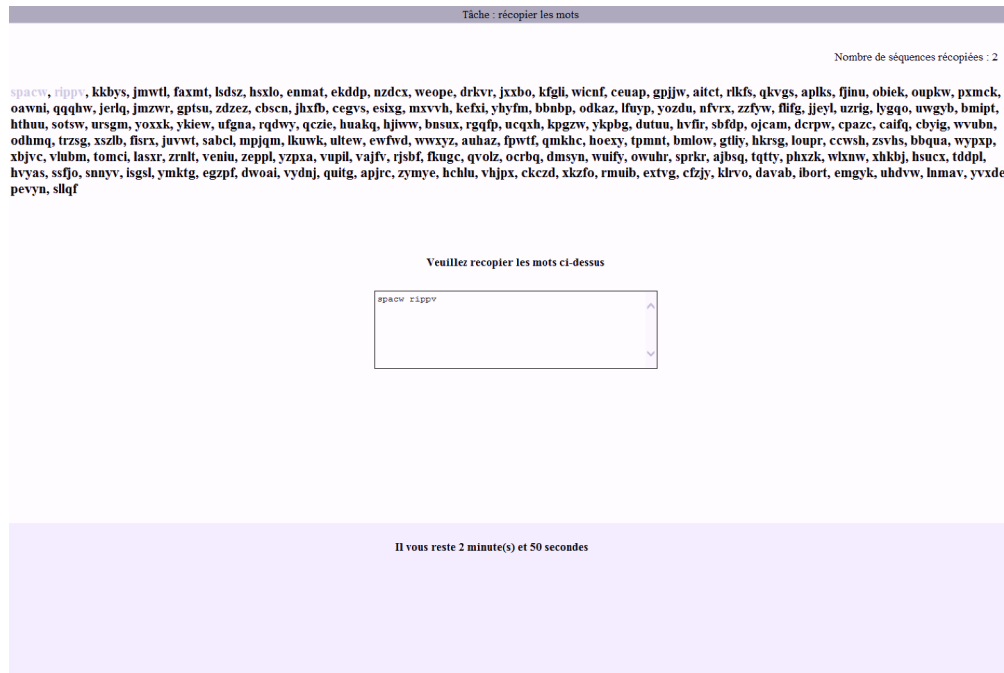


Figure C.4: Copy task.

C.1.2 Instructions

These are a rough translation of the instructions associated with the video treatment.

You are participating in an experiment where you will earn money. Your earnings will depend on your decisions and chance. They will not depend on the decisions made by other participants. We are interested in your decisions. There is no right or wrong answer to the questions you will be asked.

During the experiment, we will ask you to answer some questions that will allow us to know you better and understand the decisions you have made.

All this information is strictly anonymous and confidential.

C.1.2.1 Gains in the experiment

Your gains in the experiment are:

1. A lump-sum payment of 5€ for your participation in the experiment;
2. Plus: the remuneration corresponding to your performance in the task drawn, usually between 2 and 5€. **Correctly choosing the task you will perform can significantly increase your earnings;**
3. Plus: the gain associated with the selection of the task (between 0 and 12 cents);
4. Plus: the amount you earned in the fourth part (between 0 and 5€).

If in your earnings, hundredths of a cent appear (for example, 2.10 cents), your earnings are computed as follows:

- You win the cent amount with certainty (here 2 cents);
- The decimal amount (here 0.10) corresponds to your odds of earning an extra cent (here 10 chances out of 100).

In this example, you earn 2 cents with 90 chances out of 100 and 3 cents with 10 chances out of 100.

C.1.2.2 Remarks

The questions may look similar to you, but they are all different. Please forgive us for the sometimes slow load times. If they are extremely slow, please raise your hand.

C.1.2.3 Instructions

For a smooth experiment:

- Please put away your cell phones and put them in silent mode;
- Please do not communicate with other participants during the experiment. The whole experiment takes place on your computer;
- Calculators are forbidden during the experiment;
- If you want to ask a question, please raise your hand, someone will come to answer you.

The experience you participate in is divided into six parts. Before each part, explanations will be displayed on your screen. You will receive your payment at the end of the experiment.

C.1.2.3.1 First part

In this part, you will select the task you will perform at the end of the experiment. It is made of 11 periods. Each period is divided into two screens: a choice screen and a confirmation screen. Earnings are associated with each period. The first screen (the choice screen) allows you to select one or more of the tasks displayed (between 2 and 4), by answering “yes” or “no” to each task. You need to keep two things in mind:

- You must select at least one task;
- Each time you select a task, your earnings can increase by a few cents. The exact amount of this gain is specified on the screen;

Once you have made your selection, click on the “choices made” button. The second screen appears, with a summary of your selection in the previous screen: the tasks you have selected and the associated gain. If you are satisfied by your choice, then confirm it by clicking on the button “I confirm my choices”. If you want to change it, then click on the button “I want to change my choices”, then you will come back to the first screen. You can change your selection as many times as you like. Your selection will only be taken into account once confirmed.

At the end of the experiment, the computer will randomly draw one of the periods. You will then receive the payoff associated with this period and perform one of the tasks you have selected during this period. The task you will perform is determined as follows:

1. If you have selected more than one task, the computer will draw one at random. *For example, assuming that the gain is 10 cents per task selected and you have to choose between fill, close, or stamp envelopes and you select fill and close envelopes, you will have a one in two chance*

to perform the filling of the envelopes and on in two chances to close the envelopes. The gain related to your selection in this part is 20 cents (since you have selected two tasks) in addition to the remuneration in the task.

2. *If you have selected only one task, you will perform this task. Using the same example, if you only selected the closing of the envelopes, you will close envelopes, and the gain related to your selection in this part is 10 cents in addition to the remuneration in the task.*

Your selection and the associated gain are taken into account only once confirmed. Only one of the 11 periods in this part will be used to assign you a task and compute the associated gain. It means in particular that you will only earn one of the earnings associated with your selections and will only perform one task.

C.1.2.3.2 Tasks

You have 3 minutes to earn a maximum of money and increase your pay. You will have the choice between 4 tasks that I will present to you on the screen in a few moments. The tasks you can perform are:

Addition Additions will appear sequentially on the screen. You must give the result of these additions to increase your remuneration.

Spell-check A text in French will appear on the screen. This text contains spelling and grammar mistakes that you must correct to increase your pay.

Memory On your screen will appear successively letters, each displayed for a split second. When the sequence stops, you will be asked to give the last three letters appeared, in order, to increase your remuneration.

Copy On your screen will appear several sequences of letters separated by commas. You must copy these sequences to increase your pay.

C.2 Explanatory Power of Different Axioms

With a set of four alternatives, there are 26,254,935 possible choice correspondences and 20,736 possible choice functions. We have tested all the properties on each. Table C.1 gathers the results.¹ Most axioms have meager or low pass rates, of the magnitude of the percent or below. A random choice is not very likely to satisfy the properties. The β axiom bites a on choice correspondences and does not bite on choice functions. Lastly, on choice functions, many axioms have the same pass

¹The procedure we have used to compute the result is to consider that each choice function or choice correspondence has an equal probability of being drawn. It is *as if* the decision maker chose in every choice set a subset at random, using a uniform distribution. It is also *as if* the decision maker chose each alternative in each choice set with a coin toss.

Table C.1: Probabilities that a decision maker choosing with a uniform random satisfies a property.

Property	Choice	
	Correspondence	Function
Classical preferences	2.9e-06	1.2e-03
Semi-order	7.0e-06	1.2e-03
Interval order	7.9e-06	1.2e-03
Partial order	8.3e-06	1.2e-03
Monotone threshold	7.3e-04	1.2e-03
Menu-dependent threshold	1.4e-02	1.2e-03
Context-dependent threshold	7.3e-02	1.2e-03
α	5.7e-04	1.2e-03
β	2.6e-02	1.0e+00
Functional asymmetry	1.5e-02	1.2e-03
Jamison-Lau-Fishburn	1.5e-01	1.0e-02
Acyclicity with weak RP	1.4e-01	7.3e-01

rate and are satisfied by the same choice functions. Choice correspondences are useful because they allow us to separate the different models. The last line shows that using weak revealed preferences and acyclicity instead of strict revealed preferences weakens a lot the explanatory power when the grand set contains four alternatives, in particular on choice functions.

C.3 Size of the Chosen Sets

One of the defining characteristics of pay-for-certainty is to offer subjects the possibility to choose several alternatives. Table C.2 shows the size of the chosen sets depending on the size of the choice sets. Overall ε -correspondences, 57.64% of chosen sets are not singletons. In the 0-correspondence treatment, when subjects have no incentive to choose several alternatives, 15% of them always choose singletons. For these subjects, eliciting a choice function is not a restriction. On average, however, subjects took the opportunity to choose several alternatives. Forcing subjects to choose only one alternative is a restriction in these cases.

Table C.2 also shows that the size of the chosen sets grows relatively faster than the size of the choice sets. For instance, the ratio of chosen sets of size one over chosen sets of size two decreases when the size of the choice set increases. One explanation for this phenomenon might be that choosing from larger sets is more complicated than choosing from smaller sets, as studied in Section 4.4.

The incentive to choose more than one alternative introduced by pay-for-certainty is one driver of

Table C.2: Proportion of chosen sets of different sizes for 0-correspondences.

		Chosen set size			
		1	2	3	4
Choice set size	2	68.9%	31.1%		
	3	41.9%	40.7%	17.4%	
	4	28.3%	31.7%	30.0%	10.0%

Table C.3: Proportion of singletons chosen at different payments, depending on the size of the choice set.

		Bonus payment		
		No	Low	High
Choice set size	2	68.9%	49.6%	49.1%
	3	41.9%	21.5%	19.6%
	4	28.3%	11.7%	12.3%

Significance

Differences between no gains and positive gains are significant. Differences between low and high gains are not.

non-singleton choice, but not the only one. Table C.3 shows that indeed, the proportion of singletons chosen drops when the bonus is introduced. The introduction of the additional payment has a non-linear effect on choices: a small gain significantly changes the aggregate proportion of singletons chosen. Small variations inside the gain domains do not have many effects. There is at least two possible explanation for this phenomenon. The first is the salience of even minimal monetary gains in experiments. The second is that very few subjects have a weak strict preference. The jump between no and low gains are due to subjects who, when indifferent between two alternatives and faced with no gains, did not select all their maximal alternatives. When incentivized to choose all the maximal alternatives, they did so. The high gain is then not enough for them to bundle other alternatives with their maximal ones. We cannot disentangle the two explanations in the experiment. This aggregate behavior is compatible with partial identification but is incompatible with the full identification of the choice correspondence for all subjects.

Table C.4 shows another driver of non-singleton choice: lack of information. The proportion of singletons chosen grows when the information on the tasks is more precise. The additional information provided by the video has a substantial effect on the choice of singletons. The additional information provided by the training is not as valuable for subjects, judging by the small additional proportion of singletons chosen.

The non-singletons choice is a robust finding in the experiment. It persists even when providing better information on the alternatives or looking only at no gain of choosing more alternatives.

Table C.4: Proportion of chosen sets that are singletons, depending on the information and the size of the choice set.

Set Size	Information		
	Sentence	Video	Training
2	45.6%	56.8%	62.5%
3	17.5%	30.0%	30.3%
4	9.2%	19.5%	18.5%
N	87	328	108

C.4 Weak Revealed Preferences

In Section 4.3, we have assumed that revealed preferences are strict. That is, a chosen alternative is strictly better than an unchosen one. Strictly speaking, this assumption is not warranted with forced single choice and 0-correspondences. With forced single choice, if subjects are indifferent, we force them to select only one alternative. With 0-correspondences, as they have no incentive to choose several alternatives, they can select on their own when they are indifferent. In both cases, it implies that in theory, only saying that a chosen is weakly better than an unchosen one is right. This part is therefore dedicated to exploring weak revealed preferences, first on forced single choice, and then on choice correspondences. We use weak revealed preferences as defined in Definition 1.6. We then test the acyclicity of the preference relation, as given in Definition 1.10, as it is the necessary and sufficient condition for weak revealed preferences to be classical. The results should be read keeping in mind the last row of Table C.1: weak revealed preferences with acyclicity are more than 600 times more likely to rationalize a choice function than strict revealed preferences with acyclicity in our experiment. With choice correspondences, it is more than 4,000 times.

C.4.1 With Forced Single Choice

80% of the subjects are acyclic with choice functions and thus are rationalized with classical preferences. It is significantly higher than the 57% of subjects which can be rationalized using strict revealed preferences, the p-value of the Fisher exact test is 0.018.

For the 41 subjects whose choices can be rationalized by weak revealed preferences, we can build these preferences and explore indifference. We could not do that in Section 4.3.1. We find that 71% of the subjects have only strict preferences, while the remaining 29% have one indifference relation and five strict preferences relations. The average number of indifference relations is therefore 0.29, and it is significantly different from 0.²

²The p-value of the two-sided paired t-test of equality to 0 is lower than 0.001.

Table C.5: Proportion of each task being maximal with forced single choices, restricted to subjects who are rationalized by classical preferences.

	Task				N
	Addition	Spell-check	Memory	Copy	
Strict RP	24%	24%	21%	31%	29
Weak RP	29%	24%	29%	39%	41
P-value ^a	0.85	1	0.6	0.67	

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better.

^a P-values are given for the Fisher exact test, column-wise.

Finally, we can compare the maximal alternatives obtained with weak and strict revealed preferences. The former introduces the possibility of indifference. Table C.5 shows that copy is the most preferred alternative on average, whereas spellcheck is the least preferred.³ The proportions are not significantly different between weak and strict revealed preferences, and the social ordering would be similar.

C.4.2 0-Correspondences

85% of the subjects are acyclic with weak revealed preferences on 0-correspondences and thus are rationalized by classical preferences. It is significantly higher than 40% of subjects which can be rationalized using strict revealed preferences.⁴ It is not significantly different from the result obtained on choice functions.⁵ The acyclicity requirement with weak revealed preferences is much weaker than strict revealed preferences and much higher than with choice functions.

We build the preferences using weak revelation for the 153 subjects which satisfy the acyclicity requirement. We find that on average, each subject has 2.59 binary relations that are indifference. The average is significantly different from both 0-correspondences with strict revealed preferences and forced single choice with weak revealed preferences.⁶ The distribution is not significantly different from the 0-correspondence with strict revealed preferences, however, but it is significantly different from the force single choice elicitation.⁷ It is, however, quite heterogeneous, as Figure C.5 shows. Notably, even with weak revealed preferences, some subjects have only strict preferences.

³The sum of the proportion is higher than 100% because some subjects have several maximal alternatives when they are indifferent.

⁴The p-value of the Fisher exact test is below 0.001.

⁵The p-value of the Fisher exact test is 0.55.

⁶The p-values of two-sided two-sample t-test of equal means are lower than 0.001 in both cases.

⁷The p-value of the first KS-test is 0.06, and the second is lower than 0.001.

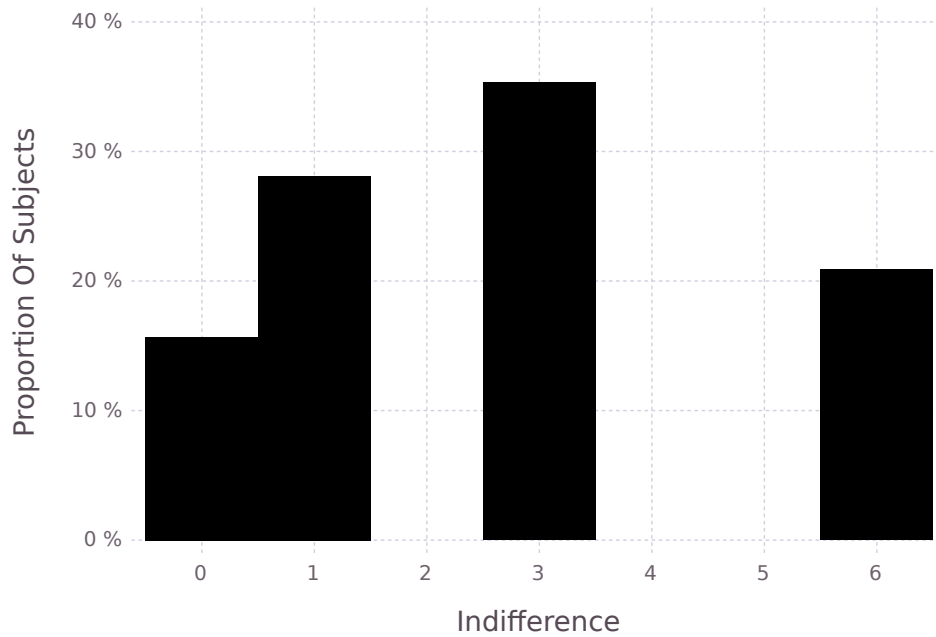


Figure C.5: Histogram of the number of indifference relations.

Finally, we can compare the maximal alternatives obtained with weak and strict revealed preferences. Table C.6 shows that copy is the most preferred alternative on average, whereas spell-check is the least preferred. The proportions are mostly not significantly different between weak and strict revealed preferences, and the social ordering would be similar.

Overall, using weak revealed preferences instead of strict revealed preferences increases significantly observed choices that can be rationalized by classical preferences. With forced single choice, it reveals a significant proportion of indifference, but it still underestimates it compared to correspondences.

C.5 Results with High Payments

In the experiment, 163 subjects faced two positive bonus payment levels: the 1 cent bonus payment level and the 12 cents bonus payment level. In the core of the paper, we reported results only for the low payment levels, as the results obtained with the high payment levels are not very different.

Classical preferences rationalize 47% of the subjects in the high payment treatment (76 subjects), which is not significantly different from the 44% in the low payment treatment.⁸ The average number of indifference relation is 3.25, which is not significantly higher than in the low payment.⁹

⁸The p-value of the Fisher exact test is 0.26.

⁹The p-value of the two-sample two-sided t-test of equality is 0.67.

Table C.6: Proportion of each task being maximal with 0-correspondences, restricted to subjects who are rationalized by classical preferences. P-values are given for the Fisher exact test, column-wise.

	Task				N
	Addition	Spell-check	Memory	Copy	
Strict RP	51%	51%	52%	58%	79
Weak RP	65%	59%	63%	70%	153
P-value	0.054	0.251	0.122	0.103	

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better.

Table C.7: Proportion of each task being maximal with 1 and 12-correspondences, restricted to subjects who are rationalized by classical preferences.

	Task				N
	Addition	Spell-check	Memory	Copy	
1-correspondence	72%	72%	65%	78%	79
12-correspondence	76%	70%	67%	78%	153
P-value ^a	0.7	0.88	0.95	1	

Note:

An alternative is maximal if, according to the revealed preference of the decision maker, no other alternative is better.

^a P-values are given for the Fisher exact test, column-wise.

Figure C.6 shows the histogram of the number of subjects with their number of indifference relations. They are quite similar, but according to a KS-test, the distributions are different.¹⁰

We can also compare the maximal alternatives obtained with 1 and 12-correspondences. Table C.7 shows that the figures are quite similar, and indeed, the differences are not significant.

If we use 0- and 12-correspondences for identification purposes instead of 0- and 1-correspondences, we fully identify the choice correspondences of 15% of the subjects and partially identify them for another 46%. Both figures are not significantly different than what we obtained in Section 4.3.4, using 0- and 1-correspondences.¹¹

Studying 1 and 12-correspondences shows that the difference in results between the two treatments is mostly insignificant. It assuages the concern that 12 cents level was too high compared to the lowest increment of 10 cents for each success in a task. This section also shows that the bonus payments have a non-linear effect, with a jump when the bonus becomes strictly positive, and then

¹⁰The p-value of the KS-test is lower than 0.001.

¹¹The p-values of the Fisher exact tests are respectively 0.55 and 0.31.

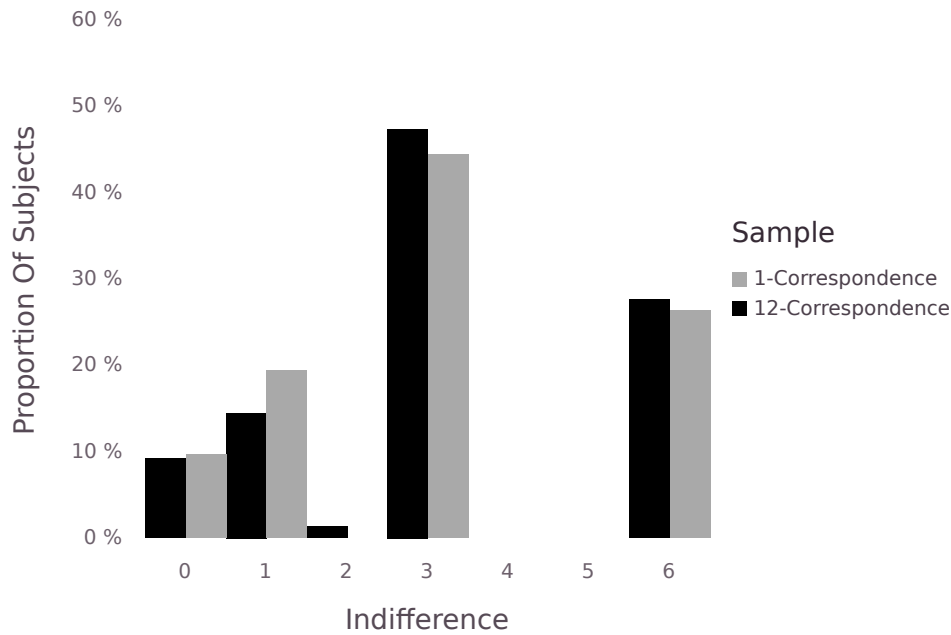


Figure C.6: Histogram of the number of indifference relations with 1- and 12-correspondences

small variation when the gains increase.

C.6 Choice Functions and Choice Correspondences on the Same Subjects

We can look at the 17 subjects for which we have both forced single choice and 0- and 1-correspondences. We use weak revealed preferences with forced single choice, in order to use most of our sample, and because it encompasses strict revealed preferences. Table C.8 lists the preferences obtained for each subject, with the different treatments. When the preferences obtained with forced single choice are only strict, the subjects satisfy WARP in addition to acyclicity with weak revealed preferences. For 0- and 1-correspondences, we use preferences obtained with the context-dependent choice model, which is a superset of classical preferences.

We can see on the Table C.8 that most of the time, choice functions correctly identify one or two of the maximal alternatives. It also fails to identify the full extent of indifference in general. We identify the choice correspondences of two subjects: 1210 and 1217. For the latter, force single choice yield consistent preferences, but it is not the case for the former.

This small subsample seems to indicate that the use of forced single choice or 0 or 1-correspondences depends on what we are seeking. If we are interested in finding one of the best alternatives and have no collective welfare purposes in mind, forced single choice looks fine. If on the other hand, we

Table C.8: Preferences deduced for each subject, depending on the elicitation method. Maximal tells us whether the alternatives that are maximal with forced single choice are also maximal according to both 0- and 1-correspondences.

Subject	Observations			WARP			Maximal
	FSC	0-C	1-C	FSC	0-C	1-C	
1201	$C \succ M \succ A \succ S$	$A \sim M \sim C \succ S$	$C \succ M \sim A \succ S$	Yes	No	No	Yes
1202	$C \succ M \succ A \succ S$	$C \succ M \succ A \succ S$	$A \sim S \sim M \sim C$	Yes	Yes	Yes	Yes
1203	$C \succ M \succ S \succ A$		$M \sim C \succ A \sim S$	Yes	No	No	Yes
1204		$M \succ C \succ A \succ S$	$A \sim M \sim C \succ S$	No	Yes	Yes	
1205	$M \succ C \succ A \succ S$	$A \sim S \sim M \sim C$	$A \sim M \sim C \succ S$	Yes	No	No	Yes
1206	$M \sim C \succ S \succ A$	$M \sim C \succ S \succ A$	$M \sim C \succ S \succ A$	No	Yes	No	Yes
1207	$A \sim C \succ S \succ M$	$A \sim S \sim C \succ M$	$A \sim C \succ S \succ M$	No	No	No	Yes
1208	$C \succ A \succ M \succ S$	$A \sim S \sim M \sim C$	$A \sim S \sim M \sim C$	Yes	No	Yes	Yes
1209	$M \succ A \succ C \succ S$	$A \sim M \succ C \succ S$	$A \succ M \succ C \succ S$	Yes	Yes	No	No
1210^a		$A \sim S \sim C \succ M$	$A \sim S \sim C \succ M$	No	Yes	Yes	No
1211	$M \sim C \succ S \succ A$	$S \sim M \succ C \succ A$	$S \sim M \succ C \succ A$	No	No	No	
1212		$S \sim C \succ M \succ A$	$S \sim M \sim C \succ A$	No	No	Yes	Yes
1213	$A \succ M \sim C \succ S$	$A \sim M \sim C \succ S$	$A \sim M \sim C \succ S$	No	No	No	Yes
1214	$S \succ M \succ C \succ A$	$A \sim S \sim M \sim C$	$S \sim M \sim C \succ A$	Yes	Yes	No	Yes
1215	$S \succ A \succ M \succ C$	$A \sim S \sim M \sim C$	$A \sim S \sim M \succ C$	Yes	Yes	No	Yes
1216	$S \sim C \succ M \succ A$		$C \sim S \succ M \succ A$	No	No	No	Yes
1217^a	$S \succ A \succ M \succ C$	$S \succ A \succ M \succ C$	$S \succ A \succ M \succ C$	Yes	Yes	Yes	Yes

Legend

A is for Addition, S for Spell-check, M for Memory and C for Copy; FSC for forced single choice, 0-C for 0-correspondence and 1-C for 1-correspondence.

^a Subjects in bold have fully identified choice correspondences with set inclusion.

want to study indifference or find a social optimum, 0 or 1-correspondence seem more appropriate.

C.7 Aggregating Choices

We look at how robust the aggregation we have proposed is. First, we introduced the information treatment. Second, we look at another determination of the best collective alternatives.

C.7.1 Information and Maximal Alternatives

Table C.9 show that the information provided only matters for the choice of the memory task on aggregate. We use the full sample of 0-, 1- and 12-correspondences, as the results of Appendix C.5 indicates that results with high gains are not significantly different from results with low gains. The

Table C.9: Regressions explaining the choice of an alternative, with and without fixed effects at the subject level.

	Task							
	Addition		Spell-check		Memory		Copy	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Intercept	0.374*** (0.098)		0.261** (0.100)		0.465*** (0.098)		0.370*** (0.098)	
Bonus	0.230 (0.111)	0.210** (0.071)	0.185 (0.113)	0.173* (0.071)	0.379*** (0.100)	0.375*** (0.099)	0.247* (0.111)	0.252** (0.084)
Video	0.041 (0.108)		0.021 (0.108)		-0.057 (0.108)		0.042 (0.108)	
Training	-0.073 (0.126)		0.157 (0.134)		-0.301* (0.120)		0.200 (0.125)	
Classical	0.182*** (0.044)	0.157*** (0.038)	0.248*** (0.044)	0.169*** (0.039)	0.153*** (0.042)	0.128** (0.041)	0.188*** (0.043)	0.292*** (0.040)
Bonus & Video	-0.109 (0.127)	-0.094 (0.078)	-0.055 (0.128)	-0.049 (0.079)	-0.299* (0.118)	-0.326** (0.106)	-0.121 (0.127)	-0.120 (0.093)
Bonus & Training	-0.092 (0.150)	-0.102 (0.101)	-0.108 (0.158)	-0.094 (0.095)	-0.343* (0.134)	-0.362** (0.122)	-0.229 (0.148)	-0.221* (0.099)
Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
N	487	483	487	483	487	483	487	483
Adjusted R ²	0.061	0.815	0.081	0.808	0.167	0.781	0.060	0.757

Note:

Robust standard errors in parenthesis. Significance: *: $p < 0.05$, **: $p < 0.01$, and ***: $p < 0.001$.

dummy variable bonus has a value of one as soon as the bonus payment is positive, i.e., with 1- and 12-correspondences. The results are not significantly different if we exclude 12-correspondences. As the information only varies between subjects, the fixed effects capture the direct effect of the information provided, which seem only to matters for the choice of the memory task in the regression without fixed effects. Subjects whose preferences can be rationalized by a classical preference exhibit more indifference, and thus have more maximal alternatives, everything else equal. They are therefore more likely to choose all alternatives, which is what we find here. The adjusted R^2 is much higher with individual fixed effects, showing that individual variations are more important than the rest to explain individual choices.

C.7.2 An Alternative Determination of Maximal Alternatives

Maximal alternatives are assessed using the preferences of the subjects. It implies that not all the sample is taken into account, as some subjects cannot be rationalized with preferences. In particular, subjects who do not satisfy WARP with forced single choice and subjects who do not satisfy fixed point with choice correspondences are ignored. In order to take them into account, we

Table C.10: Frequency of row chosen over column, in forced single choice. We can deduce the following Condorcet order: Memory \succ Copy \succ Addition \succ Spellcheck.

	Addition	Spell-check	Memory	Copy
Addition		49%	55%	57%
Spell-check	51%		51%	55%
Memory	45%	49%		41%
Copy	43%	45%	59%	

Table C.11: Frequency of row chosen from {row, column} with 0-correspondences. We can deduce the following Condorcet order: Copy \succ Memory \succ Addition \succ Spellcheck. In the choice between memory and addition, going further in the decimal shows that memory is more often chosen.

	Addition	Spell-check	Memory	Copy
Addition		61%	67%	69%
Spell-check	70%		63%	74%
Memory	67%	62%		72%
Copy	61%	58%	64%	

have to use a non-preference based approach.

One possibility to do so is to use *Condorcet winners* to determine the social ordering, instead of maximal alternatives.

Definition C.1 (winner). The alternative x is a winner in the choice between x and y if it is chosen more often than y in the choice between x and y . In case of equality, both alternatives are winners.

In other words, the frequency of $x \in c(\{x, y\})$ is higher than the frequency of $y \in c(\{x, y\})$. A *Condorcet winner* is an alternative which is a winner in all binary choices with other alternatives. To determine the social order, we then look at the Condorcet winner in all remaining alternatives and so on.

Tables C.10, C.11, C.12, and C.13 show that the order we can deduce from Condorcet winner is different between the specifications. It is strikingly different for the fully identified choice correspondences. It is also sometimes different from the order we deduce from the maximal alternatives in Table 4.1. It is notably different in the case of forced single choice, whereas it is very similar in the case of fully identified choice correspondences and identical in 0-correspondences. This difference is likely to be explained by the difference in the sample considered.

Table C.12: Frequency of row chosen from {row, column} with 1-correspondences. We can deduce the following Condorcet order: Copy \succ Memory \succ Addition \succ Spellcheck.

	Addition	Spell-check	Memory	Copy
Addition		67%	71%	78%
Spell-check	78%		73%	83%
Memory	79%	72%		81%
Copy	76%	70%	73%	

Table C.13: Frequency of row chosen from {row, column} with fully identified choice correspondences. We can deduce the following Condorcet order: Addition \sim Spellcheck \succ Copy \sim Memory.

	Addition	Spell-check	Memory	Copy
Addition		70%	67%	61%
Spell-check	70%		61%	67%
Memory	76%	70%		67%
Copy	79%	73%	67%	

Appendix D

Analytical Computer Programs

D.1 Chapter 2

Our analytical computer programs are divided into two main parts:

1. Transforming the observed data into preference relations (either RP, SUCR or TC);
2. Determining the number of cycles and the predictive power of these relations and the relationship between them. The determination of cycles is built on Johnson (1975)'s canonical algorithm.

All programs were written using the software language Julia (versions 1.0), and the standard libraries included therein. The libraries used are: CSV (0.5), DataFrames (0.18) to read and transform the original data, LightGraphs (1.2), GraphIO (0.4) to build and save the directed graph, Distributions (0.20) for the random draws, HypothesisTests (0.8) for hypothesis testing, and Gadfly (1.0), Cairo (0.6), Fontconfig (0.2) and Colors (0.9) to output plots. The regressions were run using FixedEffectsModels (0.7) and checked with Stata. The graph algorithm is a custom implementation that was added to LightGraphs afterward. The code is available [here](#). The revealed preference tests are included in a new package (under construction): RevealedPreferences.

All the codes are available from the authors, on request.

D.2 Chapter 4

The experiment was built on z-Tree (3.6.7) (Fischbacher (2007)), and subjects were recruited with Orsee (Greiner (2015)). The computer programs are similar to what was used in Chapter 2 and explained just above. They are also available on request.

Bibliography

- Adams, Abi and Ian A. Crawford (May 2015). “Models of Revealed Preference”. In: *Emerging Trends in the Social and Behavioral Sciences*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 1–15. ISBN: 9781118900772. DOI: 10.1002/9781118900772.etrds0227.
- Afriat, Sydney N. (1967). “The Construction of Utility Functions from Expenditure Data”. In: *International Economic Review* 8.1, pp. 67–77. ISSN: 00206598, 14682354. DOI: 10.2307/2525382.
- Agranov, Marina and Pietro Ortoleva (Feb. 2017). “Stochastic Choice and Preferences for Randomization”. In: *Journal of Political Economy* 125.1, pp. 40–68. ISSN: 0022-3808. DOI: 10.1086/689774.
- Aguiar, Mark and Erik Hurst (Dec. 2007). “Life-Cycle Prices and Production”. In: *American Economic Review* 97.5, pp. 1533–1559. DOI: 10.1257/aer.97.5.1533.
- Aguiar, Victor and Roberto Serrano (2017). “Slutsky matrix norms: The size, classification, and comparative statics of bounded rationality”. In: *Journal of Economic Theory* 172, pp. 163–201. ISSN: 0022-0531. DOI: <https://doi.org/10.1016/j.jet.2017.08.007>.
- Aleskerov, Fuad, Denis Bouyssou, and Bernard Monjardet (2007). *Utility maximization, choice and preference*, pp. 1–283. ISBN: 9783540341826. DOI: 10.1007/978-3-540-34183-3. arXiv: arXiv:1011.1669v3.
- Ambuehl, Sandro, B. Douglas Bernheim, and Annamaria Lusardi (Oct. 2014). *A Method for Evaluating the Quality of Financial Decision Making, with an Application to Financial Education*. Working Paper 20618. National Bureau of Economic Research. DOI: 10.3386/w20618.
- Andreoni, James, Ben Gillen, and William T. Harbaugh (2013). “The power of revealed preference tests: ex-post evaluation of experimental design”. In: *mimeo*. URL: <http://econweb.ucsd.edu/~%7B~%7Djandreoni/WorkingPapers/GARPPower.pdf>.
- Andreoni, James and Charles Sprenger (Dec. 2012). “Estimating Time Preferences from Convex Budgets”. In: *American Economic Review* 102.7, pp. 3333–56. DOI: 10.1257/aer.102.7.3333.
- Apestequia, Jose and Miguel A. Ballester (2015). “A Measure of Rationality and Welfare”. In: *Journal of Political Economy* 123.6, pp. 1278–1310. DOI: 10.1086/683838.
- Armstrong, W. E. (1939). “The Determinateness of the Utility Function”. In: *The Economic Journal* 49.195, pp. 453–467. ISSN: 00130133, 14680297. DOI: 10.2307/2224802.

- Arrow, Kenneth J. (1959). “Rational Choice Functions and Orderings”. In: *Economica* 26.102, pp. 121–127. ISSN: 00130427, 14680335. DOI: 10.2307/2550390.
- Augenblick, Ned, Muriel Niederle, and Charles Sprenger (2015). “Working Over Time: Dynamic Inconsistency in Real Effort Tasks”. In: *Quarterly Journal of Economics* 130.3, pp. 1067–1115. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjv020.Advance.
- Aumann, Robert J. (1962). “Utility Theory without the Completeness Axiom”. In: *Econometrica* 30.3, pp. 445–462. ISSN: 00129682, 14680262. DOI: 10.2307/1909888.
- Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy (Aug. 2018). “Incentives in Experiments: A Theoretical Analysis”. en. In: *Journal of Political Economy* 126.4, pp. 1472–1503. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/698136.
- Bayrak, Oben and John D. Hey (2019). “Understanding Preference Imprecision”. In: *SSRN Electronic Journal*, pp. 1–30. ISSN: 1556-5068. DOI: 10.2139/ssrn.3309108.
- Beatty, Timothy K. M. and Ian A. Crawford (Oct. 2011). “How Demanding Is the Revealed Preference Approach to Demand?” In: *American Economic Review* 101.6, pp. 2782–95. DOI: 10.1257/aer.101.6.2782.
- Benkert, Jean-Michel and Nick Netzer (2018). “Informational Requirements of Nudging”. In: *Journal of Political Economy* 126.6, pp. 2323–2355. DOI: 10.1086/700072.
- Bernheim, B. Douglas (May 2009). “Behavioral Welfare Economics”. In: *Journal of the European Economic Association* 7.2-3, pp. 267–319. ISSN: 1542-4766. DOI: 10.1162/JEEA.2009.7.2-3.267.
- (2016). “The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics”. In: *Journal of Benefit-Cost Analysis* 7.1, pp. 12–68. DOI: 10.1017/bca.2016.5.
- Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov (Sept. 2015). “The Welfare Economics of Default Options in 401(k) Plans”. In: *American Economic Review* 105.9, pp. 2798–2837. DOI: 10.1257/aer.20130907.
- Bernheim, B. Douglas and Antonio Rangel (Feb. 2009). “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics*”. In: *The Quarterly Journal of Economics* 124.1, pp. 51–104. ISSN: 0033-5533. DOI: 10.1162/qjec.2009.124.1.51.
- Bewley, Truman F. (Nov. 2002). “Knightian decision theory. Part I”. en. In: *Decisions in Economics and Finance* 25.2, pp. 79–110. ISSN: 1593-8883, 1129-6569. DOI: 10.1007/s102030200006.
- Blundell, Richard W., Martin Browning, and Ian A. Crawford (2003). “Nonparametric Engel Curves and Revealed Preference”. In: *Econometrica* 71.1, pp. 205–240. DOI: 10.1111/1468-0262.00394.
- Boccardi, Maria Jose (2018). “Predictive Ability and the Fit-Power Trade-Off in Theories of Consumer Behavior”. In: pp. 1–53. URL: <https://drive.google.com/open?id=0B-SXcE9wvakkUmNmdFY5Wm54QU0>.
- Caplin, Andrew (2016). “Economic Data Engineering”. In: May, pp. 1–44.
- Caplin, Andrew, Mark Dean, and Daniel Martin (Dec. 2011). “Search and Satisficing”. In: *American Economic Review* 101.7, pp. 2899–2922. DOI: 10.1257/aer.101.7.2899.

- Caplin, Andrew and Daniel Martin (2012). “Framing Effects and Optimization”. In: *Mimeo*. URL: <http://www.martinonline.org/daniel/CaplinMartinFramingEffects20120415.pdf>.
- Cerreia-Vioglio, Simone et al. (2018). “Deliberately Stochastic”. In: *American Economic Review*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2977320.
- Chambers, Christopher P. and Takashi Hayashi (2012). “Choice and individual welfare”. In: *Journal of Economic Theory* 147.5, pp. 1818–1849. ISSN: 0022-0531. DOI: <https://doi.org/10.1016/j.jet.2012.01.013>.
- Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman (2015). “Choice overload: A conceptual review and meta-analysis”. In: *Journal of Consumer Psychology* 25.2, pp. 333–358. ISSN: 1057-7408. DOI: 10.1016/j.jcps.2014.08.002.
- Choi, Syngjoo, Raymond Fisman, et al. (2007). “Consistency and Heterogeneity of Individual Behavior under Uncertainty”. In: *American Economic Review* 97.5, pp. 1921–1938. DOI: 10.1257/aer.97.5.1921.
- Choi, Syngjoo, Shachar Kariv, et al. (2014). “Who Is (More) Rational?” In: *American Economic Review* 104.6, pp. 1518–50. DOI: 10.1257/aer.104.6.1518.
- Costa-Gomes, Miguel, Carlos Cueva, and Georgios Gerasimou (2016). “Choice, Deferral and Consistency”. In: *School of Economics and Finance Discussion Paper No. 1416* 1416. URL: <http://www.st-andrews.ac.uk/~wwwwecon/repecfiles/4/1416.pdf>.
- (2019). “Choice, Deferral and Consistency”. In: *School of Economics and Finance Discussion Paper No. 1416*. URL: <https://www.st-andrews.ac.uk/~wwwwecon/repecfiles/4/1416.pdf>.
- Crawford, Ian A. and Bram De Rock (2014). “Empirical Revealed Preference”. In: *Annual Review of Economics* 6.1, pp. 503–524. DOI: 10.1146/annurev-economics-080213-041238.
- Danan, Eric (2008). “Revealed preference and indifferent selection”. In: *Mathematical Social Sciences* 55.1, pp. 24–37. ISSN: 0165-4896. DOI: <https://doi.org/10.1016/j.mathsocsci.2006.12.005>.
- Danan, Eric, Thibault Gajdos, and Jean-Marc Tallon (2013). “Aggregating sets of von Neumann–Morgenstern utilities”. In: *Journal of Economic Theory* 148.2, pp. 663–688. ISSN: 0022-0531. DOI: <https://doi.org/10.1016/j.jet.2012.12.018>.
- (Feb. 2015). “Harsanyi’s Aggregation Theorem with Incomplete Preferences”. In: *American Economic Journal: Microeconomics* 7.1, pp. 61–69. DOI: 10.1257/mic.20130117.
- Danan, Eric and Anthony Ziegelmeyer (2006). “Are Preferences Complete?: An Experimental Measurement of Indecisiveness Under Risk”. In: URL: <http://www.researchgate.net/profile/Anthony%20Ziegelmeyer/publication/5018150%20Are%20preferences%20complete%20An%20experimental%20measurement%20of%20indecisiveness%20under%20risk/links/0c960516d8217911fe000000.pdf>.
- De Clippel, Geoffroy and Kareen Rozen (2018). “Bounded rationality and limited datasets”. In: *mimeo*. URL: <https://www.brown.edu/Departments/Economics/Faculty/Geoffroy%20deClippel/lim-data.pdf>.

- Dean, Mark and Daniel Martin (2016). “Measuring Rationality with the Minimum Cost of Revealed Preference Violations”. In: *The Review of Economics and Statistics* 98.3, pp. 524–534. DOI: 10.1162/REST_a_00542.
- Debreu, Gérard (1954). “Representation of a Preference Ordering by a Numerical Function”. In: *Decision Processes*, pp. 159–165.
- Dhar, Ravi and Itamar Simonson (2003). “The Effect of Forced Choice on Choice”. In: *Journal of Marketing Research* 40.2, pp. 146–160. DOI: 10.1509/jmkr.40.2.146.19229.
- Dohmen, Thomas et al. (June 2011). “Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences”. In: *Journal of the European Economic Association* 9.3, pp. 522–550. ISSN: 1542-4766. DOI: 10.1111/j.1542-4774.2011.01015.x.
- Dziewulski, Pawel (2016). “Eliciting the just-noticeable difference”. In: October, pp. 1–42.
- Eliaz, Kfir and Efe A. Ok (2006). “Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences”. In: *Games and Economic Behavior* 56.1, pp. 61–86. ISSN: 0899-8256. DOI: <https://doi.org/10.1016/j.geb.2005.06.007>.
- Fischbacher, Urs (June 2007). “z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental Economics* 10.2, pp. 171–178. ISSN: 1573-6938. DOI: 10.1007/s10683-006-9159-4.
- Fishburn, Peter C. (1970). “Intransitive Indifference in Preference Theory: A Survey”. In: *Operations Research* 18.2, pp. 207–228. DOI: 10.1287/opre.18.2.207.
- Frederick, Shane, George Loewenstein, and Ted O’Donoghue (June 2002). “Time Discounting and Time Preference: A Critical Review”. In: *Journal of Economic Literature* 40.2, pp. 351–401. DOI: 10.1257/002205102320161311.
- Frick, Mira (2016). “Monotone threshold representations”. In: *Theoretical Economics* 11.3, pp. 757–772. DOI: 10.3982/TE1547.
- Gerasimou, Georgios (Aug. 2017). “Indecisiveness, Undesirability and Overload Revealed Through Rational Choice Deferral”. In: *The Economic Journal* 128.614, pp. 2450–2479. ISSN: 0013-0133. DOI: 10.1111/eoj.12500.
- Greiner, Ben (July 2015). “Subject pool recruitment procedures: organizing experiments with ORSEE”. In: *Journal of the Economic Science Association* 1.1, pp. 114–125. ISSN: 2199-6784. DOI: 10.1007/s40881-015-0004-4.
- Handbury, Jessie (2013). “Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities”. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2516748.
- Hinnosaar, Marit (2016). “Time inconsistency and alcohol sales restrictions”. In: *European Economic Review* 87, pp. 108–131. ISSN: 0014-2921. DOI: 10.1016/j.euroecorev.2016.04.012.
- Houthakker, H. S. (May 1950). “Revealed Preference and the Utility Function”. In: *Economica*. New Series 17.66, pp. 159–174. ISSN: 0013-0427. DOI: 10.2307/2549382.
- Johnson, D. (1975). “Finding All the Elementary Circuits of a Directed Graph”. In: *SIAM Journal on Computing* 4.1, pp. 77–84. DOI: 10.1137/0204007.

- Koo, Anthony Y. C. (1963). “An Empirical Test of Revealed Preference Theory”. In: *Econometrica* 31.4, p. 646. ISSN: 00129682. DOI: 10.2307/1909164.
- Landry, Peter and Ryan Webb (2017). “Pairwise Normalization: A Neuroeconomic Theory of Multi-Attribute Choice”. en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2963863.
- Little, I. M. D. (1949). “A Reformulation of the Theory of Consumer’s Behaviour”. In: *Oxford Economic Papers* 1.1, pp. 90–99. ISSN: 00307653, 14643812. DOI: 10.1093/oxfordjournals.oep.a041063.
- Luce, R. Duncan (1956). “Semiorders and a Theory of Utility Discrimination”. In: *Econometrica* 24.2, pp. 178–191. ISSN: 00129682, 14680262. DOI: 10.2307/1905751.
- Mandler, Michael (2005). “Incomplete preferences and rational intransitivity of choice”. In: *Games and Economic Behavior* 50.2, pp. 255–277. ISSN: 0899-8256. DOI: <https://doi.org/10.1016/j.geb.2004.02.007>.
- Maniquet, François and Philippe Mongin (Mar. 2015). “Approval voting and Arrow’s impossibility theorem”. en. In: *Social Choice and Welfare* 44.3, pp. 519–532. ISSN: 0176-1714, 1432-217X. DOI: 10.1007/s00355-014-0847-2.
- Manzini, Paola and Marco Mariotti (2010). “Revealed preferences and boundedly rational choice procedures: an experiment”. In: *University of St. Andrews and IZA. Working Paper. Make Us Smart., ed. Gerd Gigerenzer, Peter M. Todd and the ABC Research Group*, pp. 141–167. URL: <http://ftp.iza.org/dp2341.pdf>.
- (2012). “Categorize then Choose: Boundedly Rational Choice and Welfare”. In: *Journal of the European Economic Association* 10.5, pp. 1141–1165. DOI: 10.1111/j.1542-4774.2012.01078.x.
- (2014). “Welfare economics and bounded rationality: the case for model-based approaches”. In: *Journal of Economic Methodology* 21.4, pp. 343–360. ISSN: 1350-178X. DOI: 10.1080/1350178X.2014.965909.
- Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay (May 2012). “Revealed Attention”. In: *American Economic Review* 102.5, pp. 2183–2205. DOI: 10.1257/aer.102.5.2183.
- Mongin, Philippe and Claude d’Aspremont (1998). “Utility theory and ethics”. In: vol. 1. Kluwer Academic Publishers.
- Nehring, Klaus (June 1997). “Rational choice and revealed preference without binariness”. In: *Social Choice and Welfare* 14.3, pp. 403–425. ISSN: 1432-217X. DOI: 10.1007/s003550050075.
- Nishimura, Hiroki (2018). “The transitive core: inference of welfare from nontransitive preference relations”. In: *Theoretical Economics* 13.2, pp. 579–606. DOI: 10.3982/TE1769.
- Ok, Efe A. (2002). “Utility Representation of an Incomplete Preference Relation”. In: *Journal of Economic Theory* 104.2, pp. 429–449. ISSN: 0022-0531. DOI: <https://doi.org/10.1006/jeth.2001.2814>.
- Ok, Efe A and Gerelt Tserenjigmid (Apr. 2019). “Deterministic Rationality of Stochastic Choice Behavior”. en. In: p. 39. URL: https://drive.google.com/file/d/1K2xkmG5_0sIxiqW7PGPjP1cMAWGvfbyK/.

- Ong, Qiyan and Jianying Qiu (2018). “Identifying indecisiveness from choices”. In: 31.0.
- Pesendorfer, Wolfgang and Faruk Gul (July 2009). “A Comment on Bernheim’s Appraisal of Neuroeconomics”. In: *American Economic Journal: Microeconomics* 1, pp. 42–47. DOI: 10.1257/mic.1.2.42.
- “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019). In: *Journal of Economic Theory* 181, pp. 547–585. ISSN: 0022-0531. DOI: 10.1016/j.jet.2019.03.002.
- Richter, Marcel K. (1966). “Revealed Preference Theory”. In: *Econometrica* 34.3, pp. 635–645. ISSN: 00129682, 14680262. DOI: 10.2307/1909773.
- Rubinstein, Ariel and Yuval Salant (Jan. 2012). “Eliciting Welfare Preferences from Behavioural Data Sets”. In: *The Review of Economic Studies* 79.1, pp. 375–387. ISSN: 0034-6527. DOI: 10.1093/restud/rdr024.
- Salant, Yuval and Ariel Rubinstein (2008). “(A, f): Choice with Frames¹”. In: *The Review of Economic Studies* 75.4, pp. 1287–1296. DOI: 10.1111/j.1467-937X.2008.00510.x.
- Samuelson, Paul A (1938). “A note on the pure theory of consumer’s behaviour”. In: *Economica*, pp. 61–71. DOI: 10.2307/2548836.
- Sautua, Santiago I. (2017). “Does uncertainty cause inertia in decision making? An experimental study of the role of regret aversion and indecisiveness”. In: *Journal of Economic Behavior and Organization* 136, pp. 1–14. ISSN: 0167-2681. DOI: <https://doi.org/10.1016/j.jebo.2017.02.003>.
- Savage, Leonard J (1954). *The foundations of statistics*. Courier Corporation.
- Schwartz, Thomas (Dec. 1976). “Choice functions, “rationality” conditions, and variations on the weak axiom of revealed preference”. en. In: *Journal of Economic Theory* 13.3, pp. 414–427. ISSN: 00220531. DOI: 10.1016/0022-0531(76)90050-8.
- Selten, Reinhard (Apr. 1991). “Properties of a measure of predictive success”. In: *Mathematical Social Sciences* 21.2, pp. 153–167. ISSN: 0165-4896. DOI: 10.1016/0165-4896(91)90076-4.
- Sen, A. (July 1969). “Quasi-Transitivity, Rational Choice and Collective Decisions”. en. In: *The Review of Economic Studies* 36.3, pp. 381–393. ISSN: 0034-6527, 1467-937X. DOI: 10.2307/2296434.
- Sen, Amartya (July 1971). “Choice Functions and Revealed Preference”. In: *The Review of Economic Studies* 38.3, pp. 307–317. ISSN: 0034-6527. DOI: 10.2307/2296384.
- (1973). “Behaviour and the Concept of Preference”. In: *Economica* 40.159, pp. 241–259. ISSN: 00130427, 14680335. DOI: 10.2307/2552796.
- (1997). “Maximization and the Act of Choice”. In: *Econometrica* 65.4, pp. 745–779. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/2171939>.
- Thaler, Richard (1985). “Mental Accounting and Consumer Choice”. In: *Marketing Science* 4.3, pp. 199–214. DOI: 10.1287/mksc.4.3.199.
- Tversky, Amos and Daniel Kahneman (1974). “Judgment under Uncertainty: Heuristics and Biases”. In: *Science* 185.4157, pp. 1124–1131. ISSN: 0036-8075. DOI: 10.1126/science.185.4157.1124.

- (Jan. 1981). “The framing of decisions and the psychology of choice”. In: *Science* 211.4481, pp. 453–458. ISSN: 0036-8075. DOI: 10.1126/science.7455683.
- Tyson, Christopher J. (June 2018). “Rationalizability of menu preferences”. In: *Economic Theory* 65.4, pp. 917–934. ISSN: 1432-0479. DOI: 10.1007/s00199-017-1043-2.
- Varian, Hal R (2006). “Revealed preference”. In: *Samuelsonian economics and the twenty-first century*, pp. 99–115.
- (1982). “The Nonparametric Approach to Demand Analysis”. In: *Econometrica* 50.4, pp. 945–973. ISSN: 00129682, 14680262. DOI: 10.2307/1912771.
- Ville, Jean and P. K. Newman (1951). “The Existence-Conditions of a Total Utility Function”. In: *Review of Economic Studies* 19.2, pp. 123–128. DOI: 10.2307/2295741.

List of Tables

2.1	Installment plans	39
2.2	Average hours worked per week.	41
2.3	Income quartiles	42
2.4	Level of education	42
2.5	Average budget shares in a month	43
2.6	Percent of individuals that have cycles in RP, SUCR, and TC.	46
2.7	Average value of Selten's index by choice set size in the experimental data	52
2.8	Summary of correlations with the average Selten's index at the individual level. . .	56
2.9	Regressions of the average Selten's index for TC onto the number of direct RP cycles and the directness index (DI).	58
3.1	Fraction of the population choosing each pizza, with $0 \leq p, q \leq 1/2$	60
3.2	The observed choices of one individual, with and without the ε bonus payment. . .	76
3.3	A pair (c_0, c_ε) where the compatible preference is not unique.	79
4.1	Maximal alternatives on classical preferences.	97
4.2	Strict preferences and indifference with classical preferences.	99
4.3	Identification of choice correspondences.	100
4.4	Rationalizability by a weakening of classical preferences.	103
4.5	Each task is maximal for X% of subjects.	105
4.6	Rationalizability of 0 and 1-correspondences, depending on the information provided.	106
4.7	Average gains in the tasks.	106
4.8	Regressions of the number of indifference relations.	107

4.9	Each alternative is maximal for X% of the subjects, depending on the information.	109
A.1	Correlations with the average Selten's index at the simulation level.	115
B.1	Size of the chosen set when subjects chose their distribution over the alternatives.	119
C.1	Probabilities that a decision maker choosing with a uniform random satisfies a property.	127
C.2	Proportion of chosen sets of different sizes for 0-correspondences.	128
C.3	Proportion of singletons chosen at different payments, depending on the size of the choice set.	128
C.4	Proportion of chosen sets that are singletons, depending on the information and the size of the choice set.	129
C.5	Each task is maximal for X% of the subjects, with forced single choices and weak revealed preferences.	130
C.6	Each task is maximal for X% of the subjects, with 0-correspondences and weak revealed preferences.	132
C.7	Each task is maximal for X% of subjects, 1- and 12-correspondences.	132
C.8	Preference for each subject.	134
C.9	Regressions explaining the choice of an alternative, with and without fixed effects at the subject level.	135
C.10	Frequency of row chosen over column, in forced single choice.	136
C.11	Frequency of row chosen over column, on 0-correspondences	136
C.12	Frequency of row chosen over column, on 1-correspondences	137
C.13	Frequency of row chosen over column, on choice correspondences	137
E.1	Fraction de la population choisissant chaque pizza, avec $0 \leq p, q \leq 1/2$	163

List of Figures

2.1	Individuals in the consumption data by market.	44
2.2	Number of relation elements at the individual level, in the experimental data. . . .	48
2.3	Number of relation elements at the individual level, in the consumption data.	49
2.4	Average number of predicted alternatives for TC and SUCR at the individual level, in the experimental data.	50
2.5	Average number of predicted alternatives for TC and SUCR at the individual level, in the consumption data.	51
2.6	Average Selten's index at the individual level, in the experimental data.	53
2.7	Average Selten's index at the individual level, in the consumption data.	53
2.8	Number of direct RP cycles per individual level, in the experimental data.	54
2.9	Number of direct RP cycles per individual level, in the consumption data.	55
2.10	Directness index at the individual level, in the experimental data	57
2.11	Directness index at the individual level, in the consumption data	57
3.1	In which probability pair is each pizza the most chosen?	61
4.1	A choice screen.	93
4.2	A confirmation screen.	93
4.3	Age distribution	95
4.4	Professional occupation	96
4.5	Domain of occupation	96
4.6	Number of indifference relations.	98
4.7	Distance from partial identification	100
4.8	Number of indifference relations by information provided.	108

A.1	Size of the choice sets in the consumption data.	114
B.1	Choosing a distribution.	118
B.2	Choosing between a chosen distribution (left) and the uniform distribution (right).	118
C.1	Addition task.	122
C.2	Spell-check task.	122
C.3	Memory task.	123
C.4	Copy task.	123
C.5	Histogram of the number of indifference relations.	131
C.6	Number of indifference relations with 1- and 12-correspondences.	133
E.1	Dans quelle paire de probabilités chaque pizza est-elle la plus choisie?	164

Appendix E

Résumé en Français

E.1 Introduction

Cette thèse étudie le lien entre les choix, les préférences et le bien-être individuel. Nous allons ainsi parler tout au long de la thèse d’alternatives. Il peut s’agir d’un panier acheté au supermarché, d’un repas, de la possibilité de s’inscrire dans une université, de la date à laquelle vous dînerez avec vos parents, etc. Ces *choix* sont faits parmi différentes alternatives. Les *préférences* sont un classement des alternatives à choisir. Le *bien-être* est la satisfaction retirée d’une alternative.

L’évaluation du bien-être est une partie importante des sciences économiques, en particulier dans l’évaluation des politiques publiques et dans les recommandations politiques. Une méthode pour étudier le bien-être individuel consiste à identifier la satisfaction que les individus retiennent de leurs choix. Il peut s’agir d’une satisfaction relative ou absolue, obtenue en posant directement une question, par exemple dans le cadre d’études sur le bien-être subjectif, ou en l’inférant des préférences de chaque individu. Les préférences ne sont pas directement observables, mais elles peuvent être déduites. La méthode que nous utiliserons principalement dans cette thèse a pour origine l’article de Samuelson (1938), introduisant le concept de préférences révélées, et liant ainsi les choix et les préférences. Par la suite, Little (1949) a proposé d’utiliser les préférences révélées pour l’analyse du bien-être individuel.

Le projet de Little (1949) semble néanmoins en contradiction avec l’idée originelle de Samuelson (1938). Il affirmait ainsi que le but de son analyse était de (nous soulignons) :

Je propose donc de recommencer à zéro en attaquant directement le problème, en laissant tomber les *derniers vestiges* de l’analyse à base d’*utilité*.

Par conséquent, si les préférences révélées visaient à supprimer toute trace d’utilité dans l’analyse économique, elles ne conviendraient donc pas à l’analyse du bien-être, car elles ne représenteraient

rien de significatif de ce point de vue. Ce ne serait qu'un simple outil de modélisation des choix des individuels, une représentation mathématique commode. Comme le dit Little (1949) (la théorie du comportement du consommateur devant être comprise comme la théorie des préférences révélées) :

Dans la théorie du comportement du consommateur, telle qu'elle est formulée à l'heure actuelle, les préférences (au sens subjectif) découvertes en posant des questions sont liées au comportement du marché par le postulat selon lequel les gens essaient en réalité de maximiser la satisfaction. Il est donc faux de supposer que la théorie, telle qu'elle se présente actuellement, ne concerne que le choix. Cela concerne aussi les goûts et les aversions.

Little (1949) relie dès lors les choix à la satisfaction, mais cela ne peut être compris que si les individus ont une utilité qu'ils maximisent. Les préférences révélées tentent de déduire cet utilité à partir des choix observés. D'après Amartya Sen (1973) :

La logique de l'approche des préférences révélées réside dans cette hypothèse de révélation et non dans l'abandon de la notion de préférences sous-jacentes, malgré les dénégations occasionnels.

En résumé, l'utilisation des préférences révélées pour l'analyse du bien-être nécessite deux hypothèses. Premièrement, que les individus ont une préférence (utilité), qui reflète leur bien-être, et deuxièmement, que les préférences révélées l'identifient correctement. Pour comprendre la validité de ces hypothèses, il est utile d'expliquer le principe des préférences révélées. L'idée est assez intuitive. Si Jean choisit une pomme au lieu d'une banane lorsque les deux sont disponibles, il préfère alors probablement les pommes aux bananes. Ainsi, il aura un bien-être plus élevé avec une pomme plutôt qu'une banane. Dans cet exemple simple, nous pouvons passer directement des choix au bien-être. Les préférences sont nécessaires car tous les choix ne sont pas observés. Si nous observons que Jean choisit les pommes plutôt que les bananes et les bananes plutôt que les oranges et que l'on pense en termes de bien-être, il nous paraît logique qu'il choisisse des pommes plutôt que des oranges, même sans observer le choix entre les pommes et les oranges.

Bien qu'intuitivement simple, le raisonnement se heurte à des obstacles en pratique. Par exemple, que devrions-nous conclure si Jean choisit parfois des bananes et parfois des pommes ? Ou s'il préfère les pommes aux bananes, les bananes aux oranges et les oranges aux pommes ? Il est plus difficile de créer une préférence avec une interprétation en termes de bien-être dans ces cas. Un autre problème est que la préférence construite à partir des choix observés n'est peut-être pas celle que nous devrions utiliser pour l'analyse du bien-être. Que devrions-nous penser du choix d'un fumeur qui continue de fumer, tout en sachant qu'il met sa santé en danger, par exemple ? L'utilisation des choix comme reflétant le bien-être est l'objet de débats récurrents en économie. Amartya Sen (1973)

en a fait une première critique. Un exemple plus récent de ce débat est la discussion en économie du bien-être comportementale entre Bernheim and Rangel (2009) et Rubinstein and Salant (2012). Un exemple plus récent concerne l'économie du bien-être comportemental, entre, par exemple, Bernheim (2016) et Pesendorfer and Gul (2009). Nous reviendrons sur ce dernier débat au chapitre 2. Par ailleurs, Mongin and d'Aspremont (1998) fournissent une défense normative des préférences révélées pour l'analyse du bien-être.

Si nous supposons que nous pouvons directement déduire des choix le bien-être, au sens fort des préférences révélées, nous avons encore des problèmes pratique, comme le manque de transitivité. Koo (1963) a fourni le premier exemple, Choi, Fisman, et al. (2007), Choi, Kariv, et al. (2014), Dean and Martin (2016), entre autres fournissent des illustrations plus récentes, où il est plus difficile de déterminer une préférence à partir des choix observés, sans même évoquer le bien-être. Les préférences révélées nécessitent une certaine cohérence entre les choix observés afin de pouvoir les utiliser dans l'analyse du bien-être. En d'autres termes, les choix doivent obéir à certaines règles, de sorte que nous pouvons créer une préférence. En effet, s'il n'y a pas de règles pour les choix observés, il n'y a plus d'espoir de construire un modèle pour expliquer les choix. Le fait que, dans la pratique, certaines règles soient violées (notamment la transitivité) implique que la préférence révélée par les choix n'est pas adéquate pour l'analyse du bien-être.

Il y a deux réactions possibles à ces échecs. La première consiste à considérer que l'approche des préférences révélées est vouée à l'échec et à tenter d'apporter un éclairage sur les processus de prise de décision afin de construire un modèle de décision plus proche de la réalité. La seconde consiste à examiner les hypothèses aboutissant aux préférences révélées, et à ainsi comprendre pourquoi les préférences révélées semblent échouer. Cette thèse prend ces deux directions.

Premièrement, l'approche des préférences révélées suppose que les individus peuvent être modélisés comme s'ils maximisaient leurs préférences. Ce n'est pas ainsi que sont prises les décisions en général, comme le soulignent les recherches en psychologie et en économie. Les préférences ainsi révélées ne constituent pas un bon modèle positif des processus de décision. On peut décrire les individus comme utilisant des heuristiques pour faire des choix, un fait connu depuis au moins Tversky and Kahneman (1974). Cette compréhension a conduit au développement de l'économie comportementale. L'économie comportementale considère généralement que le choix est pertinent pour l'analyse du bien-être, même en sachant que la préférence révélée peut ne pas être la préférence sous-jacente, voire qu'elle n'existe pas.

L'étude du bien-être en économie comportementale a souvent été réalisée à l'aide de modèles *ad hoc* des heuristiques utilisées pour faire des choix, afin de déterminer la préférence quand les choix semblent incohérents. Ce type d'approches a été critiqué en raison de sa faible généralisabilité. L'économie du bien-être comportemental a émergé en réponse, à la suite des articles fondateurs de Bernheim and Rangel (2009) et de Salant and Rubinstein (2008). L'apport de ces articles a été de fournir une méthode plus rigoureuse pour l'étude du bien-être dans les situations où les choix ne

résultent pas de la maximisation de l'utilité. Jusqu'à présent, la littérature sur le sujet était essentiellement une littérature théorique (Ambuehl, Bernheim, and Lusardi (2014), Bernheim, Fradkin, and Popov (2015) sont deux exceptions). Une des contributions de cette thèse est d'apporter un éclairage empirique sur l'étude du bien-être en présence de biais. Au chapitre 2, avec mon co-auteur Daniel Martin, nous étudions deux propositions tirées de la littérature sur l'économie du bien-être comportemental.

Afin de bien positionner notre étude, il est utile de définir quelques notions. Tout d'abord, une *correspondance de choix* associe à chaque ensemble non vide d'alternative un sous-ensemble non vide de ces alternatives. Une *fonction de choix* associe à chaque ensemble non vide d'alternatives une alternative de ce même ensemble.

Les préférences révélées *fortes* sont définies de la manière suivante. Une alternative x est *révélée préférée* à une alternative y , que l'on note xR^0y , si x est choisi alors que y est disponible dans l'ensemble de choix. Une alternative x est *strictement révélée préférée* à une alternative y , que l'on note xP^0y , si x est choisi alors que y est disponible et n'est pas choisie dans l'ensemble de choix. Deux alternatives x et y sont *indifférentes*, que l'on note xI^0y , si elles sont choisies ensemble dans le même ensemble. Les préférences révélées fortes sont adaptées aux correspondances de choix, mais assez peu aux fonctions de choix.

Les préférences révélées *faibles* définissent comme précédemment R^0 , mais définissent P^0 et I^0 comme les parties antisymétrique et symétrique de la relation binaire R , c'est-à-dire que xP^0y si et seulement si xR^0y et non yR^0x , et xI^0y si et seulement si xR^0y et yR^0x . Les préférences révélées faibles sont plus adaptées aux fonctions de choix que ne le sont les préférences révélées fortes, mais elles déduisent des choix moins d'informations. Enfin, dans les deux cas, on construit les relations R , P et I comme étant les fermetures transitives de R^0 , P^0 et I^0 , respectivement.

Une préférence \succeq est dite classique si elle est réflexive, transitive et complète. Cela signifie que dans l'ensemble de toutes les alternatives, notées X , chaque élément est en relation avec lui-même (réflexive), que si $x \succeq y$ et $y \succeq z$, alors on doit avoir $x \succeq z$ (transitive) et enfin que tous les éléments doivent être en relations entre eux (complète). Les préférences classiques nous permettent de lier très facilement préférences et choix, en supposant que les individus maximisent leur préférence, ce qui est suffisant pour les ensembles finis.

La condition inverse, qui nous permet de passer des choix observés à une relation de préférence, est la condition dite d'*acyclicité*. Cette condition dit que si on observe que x est révélé strictement préféré à y (xPy), alors il est impossible que y soit révélée préférée à x (yR^0x). Cette condition n'est pas toujours vérifiée en pratique.

L'économie comportementale nous donne plusieurs explications à ces contradictions. L'idée principale est que l'individu ne maximise pas une utilité, mais au contraire utilise des heuristiques. En théorie de la décision, ces violations ont conduit à l'émergence de modèles qui relâchent les

hypothèses derrière les préférences classique. Armstrong (1939) a notamment critiqué la valeur normative de la transitivité de l'indifférence. Ces critiques sont illustrées par le fameux exemple fameux donné par Luce (1956).

On sera certainement indifférent entre une tasse de café avec 100 grains de sucres et une tasse avec 101 grains de sucres. On peut évidemment étendre cette observation à l'indifférence entre 101 et 102 grains de sucres. Cela ne signifie pas pour autant que l'on est indifférent entre un café à 100 grains de sucres et un café à 300 grains de sucre.

C'est ce que l'on appelle parfois la notion de *seuil de discrimination* : la différence entre plusieurs alternatives doit être assez élevées pour être remarquée. Luce (1956) et Fishburn (1970), entre autres, ont proposées des modèles où il faut que la différence entre deux alternatives soit assez élevée pour être remarquée. En général, ces seuils de discriminations dépendent uniquement des alternatives comparées. Aleskerov, Bouyssou, and Monjardet (2007), Frick (2016) et Tyson (2018) proposent des extensions de ces modèles en supposant que les seuils de discriminations ne dépendent pas uniquement des deux alternatives comparées, mais aussi des ensembles dans lesquelles les choix se font. Tous ces modèles nécessitent que l'on identifie des correspondances de choix pour être distingué du modèle classique. Sur les fonctions de choix, tous ces modèles sont indifférentiables.

Le chapitre 3 fournit une méthode pour étudier les préférences révélées en utilisant une méthode plus large que celle utilisée actuellement jusqu'à présent, en élargissant le domaine de choix. Il révèle certaines limites aux pratiques actuelles des préférences révélées, et notamment à la possible sous-estimation de l'indifférence avec les méthodes actuelles. Le chapitre 4 illustre à l'aide d'une expérience la méthode du chapitre E.3.

E.2 Pouvoir prédictif en économie du bien-être comportementale

Chapitre écrit avec Daniel Martin

Conserver le choix comme base des évaluations du bien-être est une idée attrayante d'un point de vue normatif. Cette idée consiste à trouver un modèle de procédures de choix, d'erreurs décisionnelles ou de biais comportementaux expliquant les choix observés, et à utiliser ce modèle pour analyser le bien-être.¹ La solution alternative, toujours basée sur les choix, consiste à générer une relation à partir de choix sans imposer beaucoup de structures *ad hoc* et d'utiliser cette relation pour effectuer une analyse de bien-être (par exemple, Bernheim and Rangel (2009), Chambers and Hayashi (2012), Apesteguia and Ballester (2015) et Nishimura (2018)).

¹Il y a de nombreux exemples en théorie de la décision et en économie comportementale, notamment Caplin, Dean, and Martin (2011), Manzini and Mariotti (2012) et Rubinstein and Salant (2012).

Ces deux approches sont très différentes, et un débat théorique a émergé sur la nécessité d'un modèle structurée pour fournir des indications précises sur le bien-être en cas de choix incohérents (Bernheim and Rangel (2009), Rubinstein and Salant (2012), Manzini and Mariotti (2014), Bernheim (2016)). Il a été avancé que les approches de comportementales du bien-être « sans modèle », qui utilisent une approche conservatrice dans la résolution des ambiguïtés normatives produites par les incohérences de choix, auront peu à dire sur le bien-être en pratique. Bien qu'il existe d'autres critères normatifs de justification des politiques publiques, si une approche à peu à dire en matière de bien-être, alors les autres implications risquent d'être sans objet.

Nous apportons des éléments empiriques à ce débat théorique en déterminant, pour des ensembles de données standards issues d'une expérience en laboratoire et de données observées, que la précision des indications de bien-être offertes par deux relations issue de l'économie comportementales du bien-être : la relation de choix stricte et sans ambiguïté (SUCR) proposée par Bernheim and Rangel (2009) et le noyau transitif (TC) proposé par Nishimura (2018). SUCR et TC sont tous deux « conservatrices », elles ne tentent pas de résoudre toutes les ambiguïtés normatives produites par les incohérences de choix. Par conséquent, leur indications sur le bien-être risque d'être incomplète lorsque les biais comportementaux ont une incidence sur le choix.

SUCR et TC ont pour objectif d'être acycliques, en excluant les éléments de relation de préférences révélés qui produisent des cycles. SUCR conserve une relation entre x et y lorsque x est strictement choisi sans ambiguïté par rapport à y (noté xP^*y), autrement dit, si et seulement si y n'est jamais choisi lorsque x et y sont disponibles.² TC conserve une relation entre x et y si toutes les comparaisons effectuées avec une troisième option z sont compatibles avec la relation entre x et y .

Nous déterminons le « pouvoir prédictif » de SUCR et TC, à savoir leur capacité à faire des prédictions précises, afin de savoir si elles offrent des indications claires sur le bien-être.³ Lorsqu'une théorie offre des prédictions multiples, le pouvoir prédictif indique à quel point ses prédictions sont précises. Étant donné que SUCR et TC peuvent être incomplètes, elles ne précisent pas toujours ce qu'un agent sélectionnerait dans un ensemble d'options, de sorte que leur pouvoir prédictif est peu clair. Par exemple, imaginez les choix de $\{x\}$ dans $\{x, y\}$, $\{x\}$ dans $\{x, y, z\}$, $\{x\}$ dans $\{x, a\}$ et $\{a\}$ dans $\{x, y, a\}$. À partir de ces choix, SUCR construit xP^*y , xP^*z et aP^*y . Pour le choix dans l'ensemble $\{x, y, z\}$, SUCR prédit que seul $\{x\}$ devrait être sélectionné. En revanche, pour les autres ensembles de choix, tels que $\{x, a\}$, SUCR prédit que n'importe quelle alternative pourrait être sélectionnée.

Le pouvoir prédictif est lié à l'évaluation du bien-être, car les prédictions d'une relation correspondent à ce qui est optimal d'après cette relation. Par exemple, si une relation de bien-être prédit qu'une seule alternative peut être sélectionnée dans un ensemble de choix, alors elle a à la fois

²Masatlioglu, Nakajima, and Ozbay (2012) a un exemple où SUCR et leur modèle donnent des indications contradictoires sur le bien-être, donc SUCR n'est pas dénuée de modèle.

³D'autres applications du pouvoir prédictif d'une relation existent, notamment par Manzini and Mariotti (2010), Beatty and Crawford (2011), Andreoni, Gillen, and Harbaugh (2013), Dean and Martin (2016) et Boccardi (2018).

un pouvoir prédictif maximal et une indication sur le bien-être précise. Toutefois, si une relation de bien-être prédit que n'importe quelle alternative peut être sélectionnée dans un ensemble de choix, elle a alors un pouvoir prédictif minimal et n'offre aucune indication en matière de bien-être. Dans l'exemple précédent, SUCR offre des indications très précises sur le bien-être dans $\{x, y, z\}$ car l'optimum de bien-être individuel pour cet ensemble de choix est x , mais elle n'offre aucune indication sur le bien-être dans $\{x, a\}$.

Une mesure naturelle de la puissance prédictive est le nombre d'alternatives prédites comme choissables, où la puissance prédictive la plus élevée possible correspond à une valeur de 1 (une seule alternative prédite) et les valeurs les plus grandes représentant une puissance prédictive inférieure. Nous utilisons l'indice de Selten (Selten (1991)) comme notre principale mesure du pouvoir prédictif. Elle est théoriquement fondée, elle a été utilisée pour des questions connexes, et elle prend aussi en compte le nombre d'options disponibles. Cet indice, axiomatisé par Selten (1991), est spécialement conçu pour les théories prédisant un sous-ensemble de résultats possibles. Avec l'indice de Selten, la proportion de choix qu'une théorie prédit avec succès au sein d'un échantillon est réduite de la « taille de la zone » prédite. Nous calculons la taille de la zone en déterminant la fraction des options devant être choisies dans un ensemble de choix.⁴

Nous testons la puissance prédictive de SUCR et de TC pour deux types de données : une expérience incitées en laboratoire, sur des ensembles de choix ; et des données « naturelles », sur des ensembles de paniers de biens. L'expérience consistait en des choix de différents plans de paiement effectués par 102 étudiants. Ces données sont issues d'une expérience réalisée par Manzini and Mariotti (2010).⁵ Les données naturelles sont composées de choix de panier de biens de 1 190 ménages constitués d'une seule personne et observée pendant 10 ans. Les données ont été récoltées par Nielsen, dans le cadre du National Consumer Panel (NCP) – anciennement connu sous le nom de Homescan Consumer Panel.⁶ Les données d'achat nous obligent à formuler des hypothèses, en particulier la séparabilité par rapport aux autres dépenses, ce qui réduit la qualité « sans modèle » de SUCR et TC. Ces hypothèses sont standard dans la littérature empirique sur les préférences révélées, et nous ne pensons pas qu'elles entravent notre analyse des validités comparées de SUCR, TC et des préférences révélées, car les relations sont obtenues avec les mêmes hypothèses.

Nous avons utilisé ces données pour quatre raisons. Premièrement, les deux sont représentatives des types de données largement utilisés dans la littérature économique. Deuxièmement, dans les deux jeux de données, les individus font des choix incohérents : dans les données expérimentales, 53% des individus font des choix qui génèrent des cycles de préférences révélées, et dans les données de consommation, 100% des individus présentent des cycles de préférences révélées.⁷ Troisième-

⁴SUCR et TC prédisent toujours qu'une alternative choisie est optimale, donc la valeur de l'indice de Selten est entièrement déterminée dans notre cas par la taille de la zone.

⁵Nous remercions les auteurs pour nous avoir fournis ces données.

⁶Données collectées par The Nielsen Company (US), LLC et fournies par le Kilts Center for Marketing Data Center de L'université de Chicago Booth School of Business.

⁷Nous restreignons la suite de l'analyse à des individus qui présentent des cycles de préférences révélées.

ment, les deux présentent des caractéristiques uniques qui les rendent assez riches pour tester efficacement le pouvoir prédictif : les données expérimentales contiennent les choix dans tous les sous-ensembles possibles, et les données de consommation contiennent un grand nombre d'individus et d'observations par individu. Quatrièmement, elles sont assez différentes les uns des autres en termes de caractéristiques démographiques, de paramètres de choix et d'alternatives.

Pour les deux jeux de données, nous trouvons que SUCR et TC ont un pouvoir prédictif élevé. Dans les données expérimentales, le nombre moyen d'alternatives optimales est 1,32 pour SUCR et 1,38 pour TC, et dans les données de consommation, il est de 1,33 pour SUCR et de 1,65 pour TC. Dans les données expérimentales, la valeur moyenne de l'indice de Selten (variant de 0 à 0,58 ici) est de 0,46 pour SUCR et de 0,44 pour TC, et dans les données de consommation, la valeur moyenne de l'indice de Selten (variant de 0 à 0,96 ici) est 0,95 pour SUCR et 0,94 pour TC.⁸

Pour savoir quand et pourquoi SUCR et TC ont un pouvoir prédictif élevé, nous étudions deux propriétés des préférences révélées (RP) qui devraient influencer leur pouvoir prédictif : le nombre de cycles directs de RP et la fraction de tous les cycles de RP qui sont directs. Dans les deux cas, nous utilisons des préférences révélées fortes.⁹ Nous disons qu'il existe un cycle de préférence révélée s'il existe x_1, x_2, \dots, x_n tels que $x_1 P x_2, \dots, x_n P x_1$ et qu'un cycle de préférence révélées est « direct » si $x P y$ et $y P x$, c'est-à-dire si sa longueur est de 2.

La longueur d'un cycle est rarement prise en compte dans l'analyse des préférences révélées au-delà de la distinction entre les violations de l'axiome faible de préférence révélée (WARP) et de l'axiome fort de préférence révélée (SARP). Néanmoins, la longueur du cycle joue ici un rôle critique.¹⁰ Par exemple, pour déterminer la puissance prédictive de SUCR et de TC, le nombre de cycles directs est plus important que le nombre de cycles de préférences de toute autre longueur et que le nombre total de cycles de préférences révélées. Le nombre de cycles de préférences révélées directs donne beaucoup d'informations sur la puissance prédictive, car SUCR et TC ne contiennent pas de relations générant des cycles directs, mais peuvent contenir des relations faisant partie de cycles de préférences révélées plus longs.

Cela dit, le nombre de cycles de RP de longueurs supérieures a de l'importance, mais ce qui compte le plus, c'est leur nombre par rapport au nombre de cycles directs de RP. S'il y a beaucoup de cycles de RP de longueur supérieure par rapport au nombre de cycles directs de RP, il s'agit d'un problème à la fois pour SUCR et TC. D'une part, SUCR sera cyclique (P^* contiendra des cycles) si des cycles de préférences révélées subsistent après la suppression de toutes les relations de RP

⁸L'indice de Selten est plus élevé dans les données de consommations en partie parce que les ensembles de choix sont plus grands en moyenne.

⁹Nous utilisons la relation de préférences révélée stricte car nous n'observons jamais plus d'une option sélectionnée dans un ensemble de choix et nous n'observons jamais deux paniers sélectionnés qui auraient pu être achetés avec la même dépense. Par ailleurs, nous n'aurions pas de cycles de longueur 2 avec des préférences révélées faible, une mesure qui nous permet de comprendre le pouvoir prédictif, comme nous allons le montrer.

¹⁰Les cycles directs de RP violent à la fois SARP et WARP, alors que les cycles de RP de longueurs supérieures ne violent que SARP.

qui sont dans des cycles directs de RP, la SUCR est donc plus susceptible de devenir cyclique s'il existe de nombreux cycles de RP de longueur supérieure par rapport au nombre de cycles directs de RP. De plus, TC exclut les relations de préférence révélés supplémentaires au-delà de ceux qui sont dans des cycles de RP directs, de sorte que sa puissance prédictive sera probablement plus faible s'il existe de nombreux cycles de RP de longueur supérieure par rapport au nombre de cycle directs de RP. Ainsi, la fraction de tous les cycles de RP qui sont directs sera importante, en plus du nombre de cycles de RP directs, pour le pouvoir prédictif du TC.

Déterminer le nombre de cycles de RP de longueur supérieure peut demander une lourde charge de calcul, nous définissons donc une limite sur la fraction de tous les cycles RP de longueur 2 en divisant le nombre de cycles de RP de longueur 2 par la somme de tous les cycles de RP de longueur 2 et 3. Nous appelons cette mesure « indice de directivité » et, à notre connaissance, cette mesure est nouvelle dans la littérature.

Nous trouvons que le nombre de cycles directs de RP est fortement et négativement corrélé avec l'indice de Selten à la fois pour SUCR et TC dans nos données. Dans les données expérimentales, la corrélation pour la SUCR est de -0,98 et le TC pour -0,92, et dans les données de consommation, la corrélation pour la SUCR est -0,75 et pour le TC est -0,81. Nous trouvons également que la fraction des cycles de préférences révélées qui sont directs est fortement et positivement corrélée à l'indice de Selten pour le TC. Dans les données expérimentales, la corrélation pour le TC est de 0,78 et dans les données de consommation, la corrélation pour le TC est de 0,69. Ces liens sont également forts, positifs et significatifs dans les régressions aux moindres carrés qui contrôlent le nombre de cycles directs de RP. En fait, telle que mesurée par R^2 , la variation de ces deux facteurs explique 89% de la variation de l'indice de Selten pour le TC dans les données expérimentales et de 67% dans les données de consommation.¹¹

En outre, nous trouvons que la fraction des cycles directs de RP est fortement et positivement corrélée à l'acyclicité de la SUCR. Dans les données de consommation, 79% des individus ont une SUCR acyclique et la corrélation est de 0,38 entre l'indice de directivité et une variable fictive prenant la valeur 1 lorsque la SUCR est acyclique.

En plus de permettre de savoir quand et pourquoi la SUCR et le TC fonctionnent bien dans la pratique, les deux mesures (le nombre de cycles de RP directs et l'indice de directivité) sont relativement rapides à calculer. Ainsi, il est facile de déterminer sur tout nouvel ensemble de données si des approches sans modèles de l'économie du bien-être comportemental sont susceptibles d'offrir une indication précise en matière de bien-être.

Cet article présente, à notre connaissance, la première application empirique non paramétrique de SUCR et la première application empirique de TC.¹² En outre, nous introduisons le pouvoir

¹¹Il n'y a pas de raisons *a priori* pour que notre indice de directivité ait de l'importance pour la SUCR, et nous n'en trouvons effectivement pas.

¹²Il existe des applications paramétriques de la SUCR, par Ambuehl, Bernheim, and Lusardi (2014) et Bernheim,

prédictif en tant qu’outil d’évaluation des relations de bien-être comportemental. Sur la base d’une mesure classique du pouvoir prédictif, nous fournissons une réponse empirique à la question du nombre d’hypothèses nécessaires afin de fournir des indications sur le bien-être. Pour les ensembles de données de choix assez standard que nous considérons, on peut donner des indications précises sur le bien-être sans faire de nombreuses hypothèses – sur la forme de l’utilité, sur la nature des biais comportementaux ou sur les ensembles de choix à prendre en compte.

E.3 Identification des correspondances de choix

Dans un article fondateur, Samuelson (1938) a introduit la méthode des préférences révélées. Il lie les préférences et les choix en supposant que les alternatives choisies sont meilleures que les alternatives non choisies, révélant ainsi la préférence d’un individu. Dans les expériences en laboratoire et dans la plupart des situations réelles, nous identifions une fonction de choix en utilisant une méthode de choix contraint : l’individu choisit exactement une alternative dans l’ensemble de choix. En théorie, en revanche, les choix sont généralement modélisés avec une correspondance de choix : l’individu choisit un ensemble non vide dans l’ensemble de choix.

Dans ce chapitre, nous introduisons *payer pour la certitude*, une méthode permettant d’identifier les correspondances de choix dans des expériences incitées. Nous autorisons les individus à choisir plusieurs alternatives, et leur donnons une petite incitation à choisir plusieurs alternatives, puis nous en sélectionnons une pour le paiement. Certaines expériences ont déjà permis aux individus de choisir plusieurs alternatives, mais aucune d’entre elles n’a tenté d’obtenir la correspondance de choix des individus d’une manière incitée. Danan and Zieglmeyer (2006) ont permis aux individus de reporter leur choix, moyennant un coût. Agranov and Ortoleva (2017) ont fait en sorte que les sujets répètent sciemment leur choix et leur ont permis de changer entre chaque choix dans le même ensemble. Ong and Qiu (2018), Sautua (2017) et “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019) ont permis aux sujets de déléguer leur choix à un dispositif aléatoire, avec et sans frais. Enfin, Costa-Gomes, Cueva, and Gerasimou (2019) ont permis aux individus de reporter leur choix, moyennant un coût. Toutes ces méthodes ne font pas du choix de toutes les alternatives maximales la stratégie dominante. C’est au mieux une stratégie faiblement dominante. Ce faisant, elles risquent de ne pas identifier la correspondance de choix exacte des individus s’ils sont indifférents entre deux alternatives ou plus.

Identifier la correspondance de choix est important pour l’analyse du bien-être. Danan, Gajdos, and Tallon (2013) et Danan, Gajdos, and Tallon (2015) ont montré que l’agrégation de préférences incomplètes pourrait être relativement plus facile que l’agrégation de préférences complètes. Maniquet and Mongin (2015) ont montré qu’il est possible d’agréger des préférences avec deux classes

d'indifférence au maximum, alors qu'avec un nombre de classes d'indifférence non limité, c'est impossible, d'après le théorème d'Arrow. Afin d'étudier l'indifférence ou les préférences incomplètes, nous avons besoin d'une correspondance à choix, comme le montrent Mandler (2005), Eliaz and Efe A. Ok (2006) et Aleskerov, Bouyssou, and Monjardet (2007). L'exemple suivant illustre pourquoi il est souhaitable d'identifier une correspondance plutôt qu'une fonction de choix.

Exemple E.1 (Pizzeria). Sur une petite île, il y a un pizzaiolo qui produit trois types de pizzas, une pizza végétarienne (V), une pizza quatre fromages (F) et une pizza au jambon (J). Il est le seul pizzaiolo de l'île, il n'y a donc pas de forces concurrentielles sur le marché des pizzas. Les préférences des insulaires, que le pizzaiolo ignore, sont les suivantes :

- La moitié des habitants de l'île sont indifférents entre la pizza végétarienne et la pizza quatre fromages et préfère les deux premières à la pizza au jambon, c'est-à-dire $V \sim F \succ J$;
- Une autre moitié est indifférente entre la pizza au jambon et la pizza quatre fromages et préfère les deux premières à la végétarienne, à savoir $J \sim F \succ V$.

Pour des raisons de coût, le pizzaiolo souhaite produire un seul type de pizza. Le choix naturel à faire dans sa situation est de conserver la pizza la plus choisie. Est-ce le meilleur choix pour le bien-être de tous les insulaires ?

Supposons que parmi la première moitié de la population, une proportion p choisit la pizza végétarienne, et que dans la deuxième moitié, une proportion q choisit la pizza au jambon. Le résumé de la probabilité que chaque pizza soit choisie est présenté dans le tableau E.1.

Table E.1: Fraction de la population choisissant chaque pizza, avec $0 \leq p, q \leq 1/2$.

	végétarienne	jambon	4 fromages
Choix observé	p	q	$1 - p - q$
Alternatives maximales	$\frac{1}{2}$	$\frac{1}{2}$	1

La figure E.1 montre les valeurs de (p, q) pour lesquelles chaque type de pizza serait choisie. Dans les deux tiers des cas, le pizzaiolo conserverait la pizza quatre fromages, alors que dans un tiers des cas, il conserverait une autre pizza, ce qui réduirait le bien-être de la population. Si les habitants de l'île tirent aléatoirement pour décider quand ils sont indifférents et que le nombre d'habitants de l'île achetant une pizza est suffisamment important, il conservera la pizza quatre fromages. En revanche, si le nombre d'observations est assez faible, il pourrait conserver les pizzas végétariennes ou au jambon. Si le pizzaiolo connaissaient les préférences des citoyens de l'île, il aurait conservé la pizza quatre fromages dans tous les cas. C'est la pizza qui ne diminue pas le bien-être des insulaires, et qu'il est donc préférable de garder du point de vue collectif.

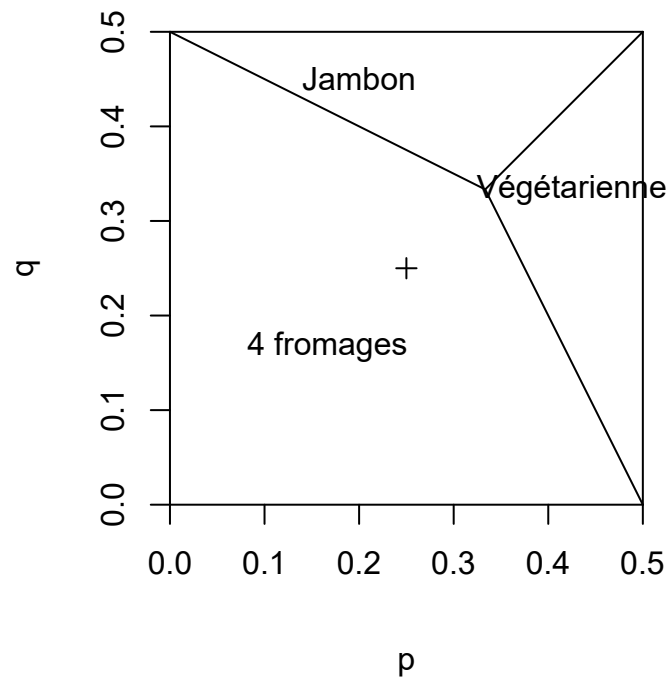


Figure E.1: Dans quelle paire de probabilités chaque pizza est-elle la plus choisie?

Cet exemple montre qu'en général, la révélation des préférences en vue d'une étude du bien-être exige davantage que l'identification d'une fonction de choix. Cela nécessite l'ensemble des alternatives maximales, c'est-à-dire la correspondance de choix pour tous les individus. Identifier une correspondance de choix pour chaque individu nous permet en outre de répondre à de nouvelles questions. Premièrement, cela nous permet d'aller au-delà du paradigme classique des préférences révélées. Deuxièmement, cela nous permet d'étudier toute l'étendue de l'indifférence.

Les correspondances de choix ne peuvent pas être obtenues directement et simplement. Une méthode simple qui permettrait de les obtenir serait la suivante : permettre aux individus de choisir plusieurs alternatives, puis en sélectionner une au hasard pour leur rémunération. Cette méthode peut échouer à identifier la correspondance de choix en présence d'indifférence. En effet, choisir toutes les alternatives maximales n'est pas une stratégie strictement dominante dans ce contexte.

Dans l'exemple E.1, imaginons maintenant que le pizzaiolo permette à tous ses clients de choisir autant de pizzas qu'ils le souhaitent. S'ils choisissent plus d'une pizza, il leur dit qu'ils auront l'une de leurs pizzas, choisie par lui au hasard. Que choisiront les clients avec les préférences $V \sim F \succ J$? Ils peuvent choisir $\{V, F\}$, $\{F\}$ ou $\{V\}$ et en retirer le même bien-être. Choisir toutes les alternatives maximales n'est pas une stratégie strictement dominante, et cela crée un problème d'unicité pour l'observateur – ici le pizzaiolo. Cette procédure ne garantit pas que l'individu choisit toutes les alternatives maximales, Pareto-supérieures. En revanche, les modèles de préférences révélées construits à partir des correspondances de choix (par exemple, Amartya Sen (1971), Schwartz (1976), Nehring (1997)) supposent que l'ensemble choisi est l'ensemble des alternatives maximales. Les modèles qui étudient la relaxation du paradigme classique en particulier reposent sur cette interprétation des correspondances de choix.

Payer pour la certitude résout le problème de l'unicité en incitant les individus à choisir toutes les alternatives maximales. Pour chaque alternative choisie, l'individu gagne un supplément $\varepsilon > 0$.¹³ Dans l'exemple précédent, il est préférable pour les clients de choisir $\{V, F\}$ et d'obtenir la pizza végétarienne ou la pizza quatre fromages et un gain supplémentaire de 2ε plutôt que de choisir et d'obtenir $\{V\}$ ou $\{C\}$ et ε . La description complète de payer pour la certitude est la suivante :

Dans chaque ensemble d'alternatives S , l'individu choisit toutes les alternatives qu'il souhaite, c'est-à-dire un sous-ensemble $c(S)$. Il gagne un (petit) bonus de $\frac{|c(S)|}{|S|}\varepsilon$ selon les alternatives choisies. L'alternative qu'il obtient est sélectionnée parmi les alternatives choisies à l'aide d'un mécanisme de sélection, par exemple un tirage au sort équiprobable.

Le supplément de payer pour la certitude implique des pertes de gains hypothétiques lorsque l'ensemble choisi n'est pas l'ensemble total, ce qui explique son nom. Le choix d'une seule alternative rapporte $\frac{1}{|S|}\varepsilon$, alors que le choix de deux rapporte $\frac{2}{|S|}\varepsilon$, et ainsi de suite. Nous montrons que,

¹³En principe monétaire, mais pas nécessairement. Tout ce qui représente un gain ou un coût pour l'individu convient.

sous certaines hypothèses de monotonie, l'individu choisit toutes les alternatives maximales. Les choix observés avec payer pour la certitude et un supplément d' ε sont appelés la correspondance- ε à partir de maintenant.

Un problème potentiel est que certaines alternatives choisies pourraient ne pas être maximales si ε est grand et que les différences entre la satisfaction retirées de certaines alternatives sont proches, à hauteur de moins de $\frac{1}{|S|}\varepsilon$ l'une de l'autre. Payer pour la certitude à de nombreux traits en communs avec l'expérience de Costa-Gomes, Cueva, and Gerasimou (2016), mais il existe quelques différences essentielles.¹⁴ Premièrement, le choix entre plusieurs alternatives implique un gain et non une perte. Deuxièmement, le gain sera beaucoup moins important et suscitera théoriquement plus de relations d'indifférence. Enfin, la stratégie dominante d'un individu indifférent entre différentes alternatives consiste à toutes les choisir.

Nous fournissons deux résultats d'identification dans ce chapitre. Le premier est un résultat d'identification à la limite, montrant que si ε est suffisamment petit et que l'ensemble de toutes les alternatives est fini, nous identifions la correspondance de choix de l'individu. Le deuxième résultat est un résultat d'identification par inclusion. Dans n'importe quel ensemble de choix, la correspondance de choix doit être comprises entre la correspondance-0 et la correspondance- ε . Lorsque ce critère est satisfait, nous disons que la correspondance de choix d'un individu est *partiellement* identifiée. L'identification est précise lorsque dans un ensemble d'alternatives données, le choix avec et sans paiement additionnel sont égaux. Lorsque l'identification est précise dans tous les ensembles de choix possible, l'identification est *complète*. La correspondance de choix complètement identifiée d'un individu est sa correspondance de choix théorique.

En pratique, le deuxième résultat est directement vérifiable, alors que le premier ne l'est pas. En général, il est impossible de garantir une identification complète pour tout le monde. Nous fournissons donc des résultats pour la compatibilité de différents modèles de prise de décision lorsque la correspondance de choix est partiellement identifiée. Ces méthodes sont constructives, elles nous permettent d'obtenir la préférence correspondante.

E.4 Payer pour la certitude dans une expérience

Ce chapitre est consacré à une expérience que nous avons menée en utilisant la méthode de paiement pour la certitude avec le mécanisme de sélection uniforme. Il a deux objectifs principaux.

Premièrement, il illustre la méthode décrite au chapitre précédent en pratique.

De l'ensemble de choix S , chaque alternative choisie ajoute un supplément de $\frac{1}{|S|}\varepsilon >$

¹⁴L'expérience utilise la théorie de Gerasimou (2017) afin d'identifier l'indécision, en utilisant en particulier des correspondances de choix qui peuvent être vide. L'objectif est différent de l'objectif de notre méthode, ce qui explique les différences d'approches.

0 par alternative au gain du sujet. Le total des suppléments versés est $\frac{|c_\varepsilon(S)|}{|S|}\varepsilon$ par ensemble de choix. L'alternative qu'il obtient est sélectionnée à l'aide d'un tirage au sort équiprobable sur l'ensemble des alternatives choisies $c_\varepsilon(S)$.

Cette expérience a trois caractéristiques principales afin d'exploiter les résultats d'identification du chapitre E.3. Premièrement, la méthode expérimentale classique, les sujets doivent y choisir une et une seule alternative dans chaque ensemble de choix. Nous appellerons cette méthode de la *fonction de choix*. Deuxièmement, payer pour la certitude avec un paiement nul, appelée à partir de maintenant la *correspondance-0*. C'est l'extension la plus simple de la méthode classique permettant des choix multiples. Troisièmement, payer pour la certitude avec un supplément total ε de 1 centime, appelé à partir de maintenant la *correspondance-1*. Nous utilisons ensuite les résultats d'identification à la limite et par inclusion prouvés au chapitre E.3. Ce dernier est plus complexe à mettre en œuvre que le premier, car il s'appuie à la fois sur la correspondance-0 et sur la correspondance-1 pour chaque sujet, alors que le premier ne s'appuie sur la correspondance-1. En échange, il nous garantit d'identifier la correspondance de choix des individus.

Deuxièmement, l'expérience a pour objectif de mieux comprendre comment modéliser les décisions individuelles. Effectuer une expérience où des ensembles sont choisis, plutôt que des alternatives, nous permet d'explorer des modèles allant au-delà des préférences classiques, et en particulier des modèles d'indifférence intransitive et de choix dépendant de l'ensemble. En excluant les expériences sur le vote par approbation, très peu d'expériences incitées ont permis aux individus de choisir plusieurs alternatives. À notre connaissance, cinq l'ont déjà fait : Danan and Ziegelmeyer (2006), Costa-Gomes, Cueva, and Gerasimou (2016), “Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization” (2019), Ong and Qiu (2018), Agranov and Ortoleva (2017), alors que de nombreux modèles de théorie de la décision ont besoin du choix d'un ensemble, entre autres, Arrow (1959), Amartya Sen (1971), Schwartz (1976), Eliaz and Efe A. Ok (2006), Aleskerov, Bouyssou, and Monjardet (2007).

Obtenir un ensemble plutôt qu'une alternative choisie permet d'étudier l'indifférence directement. En effet, en utilisant des préférences révélées faibles, les fonctions de choix donnent seulement une limite inférieure à l'indifférence. Les correspondances de choix quant à elles la quantifient exactement. Cela pourrait avoir d'importantes répercussions sur l'étude du bien-être individuel et collectif. Le théorème d'impossibilité d'Arrow montre qu'il est impossible d'agrèger les préférences des individus de manière non dictatoriale. Il est toutefois évident que, dans le cas extrême où tous les individus sont indifférents entre toutes les alternatives, il n'y aura pas de théorème d'impossibilité. Maniquet and Mongin (2015) montrent un résultat plus fort : lorsque les préférences sont divisées au maximum en deux classes d'indifférence, il est possible d'agrèger les préférences.

Dans notre expérience, les sujets ont choisi entre quatre tâches rémunérées : une tâche d'addition, une tâche de vérification orthographique, une tâche de mémorisation et une tâche de copie.

L'utilisation de tâches en économie expérimentale n'est pas nouvelle. Par exemple, Augenblick, Niederle, and Sprenger (2015) ont utilisé des tâches d'effort pour mesurer l'incohérence temporelle. Les tâches présentent plusieurs avantages pour cette expérience. Premièrement, ces tâches ont été créées pour cette expérience et les sujets n'en savaient donc rien avant la session, nous permettant ainsi de contrôler l'information qui leur est délivrée. Il y avait trois niveaux d'information, un assez faible, un où les tâches étaient présentée avec une vidéo et un dernier où ils pouvaient s'entraîner sur chaque tâche avant de les choisir. Deuxièmement, nous voulions créer un environnement où l'indifférence n'émerge pas facilement, et les préférences sont susceptibles d'être plus précises sur des tâches d'effort qu'entre des loteries, comme le montrent Augenblick, Niederle, and Sprenger (2015), par exemple, certains sujets peuvent détester le calcul et rejetteraient donc fortement la tâche d'addition. Enfin, il n'y a pas de classement objectif des tâches, et elles sont très différentes. Cela devrait fournir un ensemble de préférences relativement hétérogène, ce qui est intéressant pour étudier l'agrégation des préférences.

Afin de falsifier les différents modèles testés, tous les sujets ont été confrontés à des choix dans tous les sous-ensembles possibles des quatre tâches, et ont donc choisi au moins 11 fois. Selon la théorie du chapitre E.3, nous devons observer des choix sans gain (0 centime) et des choix avec des gains positifs et faibles par rapport aux avantages des alternatives ($\varepsilon > 0$). Nous sommes allés un peu plus loin, et la plupart des sujets ont fait face à trois niveaux différents suppléments, à savoir trois ε différents : un sans supplément (0 centime), un supplément faible (1 centime), et un supplément élevé (12 centimes). Cela nous aide à caractériser la validité du résultat d'identification à la limite démontré dans la proposition 3.2. Le supplément maximal n'est jamais supérieur à 6 centimes, ce qui est inférieur à l'incrément le plus bas pour un succès supplémentaire dans n'importe quelle tâche, qui est de 10 centimes. Nous avons également forcé certains sujets à choisir une seule alternative. Pour certains de ces sujets, nous avons donc à la fois une fonction de choix et les correspondances 0 et 1. Les résultats avec des gains élevés ne sont pas significativement différents des résultats obtenus avec des gains faibles. Nous ne présentons donc que les résultats avec des gains nuls et faibles au cœur de la thèse.

Les résultats ont été obtenus avec 214 participants au cours de 13 sessions. Nous observons que les participants choisissent plusieurs alternatives quand ils en ont la possibilité. Dans la correspondance 0, où il n'y a aucune incitation à choisir plusieurs alternatives, seuls 55,40% des choix effectués sont des alternatives uniques et seulement 15% des sujets ont toujours choisi des alternatives uniques. Nous identifions complètement la correspondance de choix de 18% des participants, et nous l'identifions partiellement pour 40% supplémentaires.

Les préférences classiques rationalisent 57% des fonctions de choix, ce que nous pourrions espérer obtenir avec la méthode expérimentale classique. En supposant que 1 centime soit suffisamment faible pour utiliser le résultat d'identification à la limite, 40% des correspondances de choix peuvent être rationalisées par des préférences classiques. Sinon, en utilisant l'identification par inclusion,

les préférences classiques rationalisent 97% des correspondances de choix complètement identifiées et sont compatibles avec 92% des correspondances de choix partiellement identifiées. La différence entre les fonctions de choix et les correspondances de choix identifiées est toujours significative.

Les préférences classiques rationalisent ou sont compatibles avec la plupart des correspondances de choix totalement ou partiellement identifiées, mais pas avec toutes les correspondances 1. Dans ce dernier cas, nous explorons des extensions des préférences classiques. Les modèles d'indifférence intransitives ne rationalisent que marginalement plus de correspondances 1 (42%). L'introduction de la dépendance du choix à l'ensemble augmente la rationalisabilité de l'échantillon. Le modèle d'optimisation occasionnelle, du à Frick (2016), rationalise 57% des correspondances 1, le modèle de choix dépendant du menu 91% et le modèle de choix dépendant du contexte 96%, tous les deux dus à Aleskerov, Bouyssou, and Monjardet (2007). Tous ces pourcentages sont significativement plus élevés que ceux obtenus avec les préférences classiques.¹⁵

Nous explorons ensuite le type de préférences que les sujets expriment lorsque leurs préférences sont classiques. Nous utilisons des préférences révélées strictes pour les fonctions de choix, on n'élucide donc pas d'indifférence. Il y a 4 tâches dans l'expérience, donc une relation de préférence complète est composée de 6 relations binaires. Avec des fonctions de choix, ces 6 relations binaires sont des préférences strictes. Lorsque nous identifions complètement les correspondances de choix, le nombre de relations binaires qui sont des indifférences est en moyenne de 2,12 par sujet. Lorsque nous utilisons la correspondance de choix partiellement identifiée, la moyenne est à 2,08. Avec les correspondances 1, la moyenne est de 3,11. Quand les participants peuvent choisir plusieurs alternatives, on observe une part d'indifférence significative, qui n'est pas capturée par les fonctions de choix.¹⁶

Enfin, nous montrons que lorsque les sujets ont plus d'informations, leurs choix sont plus précis, c'est-à-dire que la taille des ensembles choisis diminue, ainsi que leur nombre de relations d'indifférence. En effet, on pourrait penser que si les individus disposent de peu d'informations sur deux tâches, il leur est difficile de les comparer. Dans ce cas, ils peuvent parfois choisir l'un, parfois l'autre. C'est précisément le type de comportement que les modèles proposer par Aleskerov, Bouyssou, and Monjardet (2007), Frick (2016) et Tyson (2018) cherchent à expliquer.

E.5 Conclusion

Cette thèse examine le lien entre les choix, les préférences et le bien-être individuel. Nous voulions étendre le paradigme classique des préférences révélées, en allant au-delà de la rationalisabilité avec les préférences classiques, tout en conservant une compréhension relativement traditionnelle

¹⁵Utilisation d'un test exact de Fisher.

¹⁶Et ce, même si on utilise des préférences faibles, comme montré dans l'annexe C.4.

des préférences révélées. Dans le chapitre E.2, avec Daniel Martin, nous avons montré l'intérêt de nouvelles approches issues de l'économie du bien-être comportemental dans la pratique. Ces approches nous permettent de construire des préférences révélées plus robustes, à des fins d'analyse du bien-être.

Dans le chapitre E.3, nous avons construit et caractérisé une méthode qui comble une différence fondamentale entre la littérature empirique et théorique sur les préférences révélées. Nous caractérisons une méthode pour identifier les correspondances de choix en pratique. Cela nous permet d'adopter une compréhension moins restrictive de la rationalité dans la pratique, au chapitre E.4.

Les résultats présentés dans cette thèse laissent des questions ouvertes. La première question est la validité des préférences révélées aux chapitres E.2 et E.4 pour l'analyse du bien-être. Nous ne savons pas, et il est difficile de savoir, si les préférences révélées révèlent bien le bien-être des individus.

De plus, chaque chapitre laisse des questions ouvertes. La validité externe des résultats obtenus au chapitre E.4 n'est pas évidente. Les tâches ont été spécialement conçues pour cette expérience. Une expérience avec des alternatives plus communes serait intéressante et pourrait confirmer ou infirmer les résultats obtenus ici. Une des faiblesses de la méthode présentée au chapitre E.3 est son incompatibilité avec les loteries. Cela signifie qu'un des objets de choix les plus utilisés dans les expériences ne peut pas être utilisé avec payer pour la certitude.

Résumé

Les préférences révélées lient choix, préférences et bien-être lorsque les choix apparaissent cohérents. Le premier chapitre s'intéresse à la force des hypothèses nécessaire pour obtenir des indications précises sur le bien-être quand les choix sont incohérents. Il utilise les données d'expériences en laboratoire et sur le terrain pour évaluer le pouvoir prédictif de deux approches utilisant peu d'hypothèses. Ces approches ont un pouvoir prédictif élevé pour une majorité d'individus, elles fournissent donc des indications précises sur le bien-être. Le pouvoir prédictif de ces approches est fortement corrélé à deux propriétés des préférences révélées. Le deuxième chapitre introduit une méthode pour obtenir l'ensemble des meilleures alternatives d'un individu, en cohérence avec théorie des préférences révélées, mais en contradiction avec les pratiques expérimentales. Les individus sont incités à choisir plusieurs alternatives à l'aide d'un petit paiement additionnel, mais sont rémunérés à la fin par une seule, tirée au hasard. Les conditions pour que les meilleures alternatives soient partiellement ou complètement identifiées sont données. Le troisième chapitre applique cette méthode dans une expérience. Les meilleures alternatives sont complètement identifiées pour 18% des sujets et partiellement pour 40%. Les préférences complètes, réflexives et transitives rationalisent 40% des choix observés dans l'expérience. Permettre que les choix dépendant de l'ensemble de choix, tout en conservant les préférences classiques, rationalise 96% des choix observés. Enfin, on observe une quantité significative d'indifférence, bien supérieure à ce qui est obtenu traditionnellement.

Mots-clés: Préférences révélées, économie comportementale, bien-être, correspondance de choix, expérience, rationalisabilité, seuil de détermination, axiome des préférences révélées faibles

Summary

Revealed preferences link choices, preferences, and welfare when choices appear consistent. The first chapter assesses how much structure is necessary to impose on a model to provide precise welfare guidance based on *inconsistent* choices. We use data sets from the lab and field to evaluate the predictive power of two conservative “model-free” approaches of behavioral welfare analysis. We find that for most individuals, these approaches have high predictive power, which means there is little ambiguity about what should be selected from each choice set. We show that the predictive power of these approaches correlates highly with two properties of revealed preferences. The second chapter introduces a method for eliciting the *set* of best alternatives of decision makers, in line with the theory on revealed preferences, but at odds with the current practice. We allow decision makers to choose several alternatives, provide an incentive for each alternative chosen, and then randomly select one for payment. We derive the conditions under which we partially or fully identify the set of best alternatives. The third chapter applies the method in an experiment. We fully identify the set of best alternatives for 18% of subjects and partially identify it for another 40%. We show that complete, reflexive, and transitive preferences rationalize 40% of observed choices in the experiment. Going beyond, we show that allowing for menu-dependent choices while keeping classical preferences rationalize 96% of observed choices. Besides, eliciting sets allows us to conclude that indifference is significant in the experiment, and underestimate by the classical method.

Keywords Revealed preferences, behavioral economics, welfare, choice correspondences, experiment, rationalizability, just-noticeable-difference, weak axiom of revealed preferences