Eliciting Choice Correspondences

A General Method and an Experimental Implementation

*Elias Bouacida**

November, 2018

Abstract

I introduce a general method for identifying *choice correspondences* experimentally, i.e., the *sets* of best alternatives of decision makers in each choice sets. Most of the revealed preference literature assumes that decision makers can choose sets. In contrast, most experiments force the choice of a *single* alternative in each choice set. In this paper, I allow decision makers to choose several alternatives, provide a small incentive for each alternative chosen, and then randomly select one for payment. I derive the conditions under which the method at least partially identifies the choice correspondence, by obtaining supersets and subsets for each choice set. I illustrate the method with an experiment, in which subjects chose between four paid tasks. I can retrieve the full choice correspondence for 26% of subjects and bind it for another 46%. Subjects chose sets of size 2 or larger 60% of the time, whereas only 3% of them always chose singletons. I then show that 46% of all observed choices can be rationalized by complete, reflexive and transitive preferences in my experiment, i.e., satisfy the Weak Axiom of Revealed Preferences – WARP hereafter. Weakening the classical model, incomplete preferences or just-noticeable difference preferences do not rationalize more choice correspondences. Going beyond WARP, however, I show that complete, reflexive and transitive preferences with menu-dependent choices rationalize 93% of observed choices. Having elicited choice correspondences allows me to conclude that indifference is widespread in the experiment. These results pave the way for exploring various behavioral models with a unified method.

JEL Codes: C91, D01, D60

keywords: choice correspondences, revealed preferences, welfare, indifference, WARP, just-noticeable preferences, aggregation of preferences.

*Paris School of Economics, University Paris 1 Panthéon Sorbonne. elias.bouacida@psemail.eu I am grateful to Jean-Marc Tallon, Daniel Martin and Stéphane Zuber for their invaluable advice and guidance. I want to thank Maxim Frolov for his help during the experiment. I am also grateful to all the people I have interacted with during the course of my Ph.D., among them Eric Danan, Peter Klibanoff, Marco Mariotti, Pietro Ortoleva, Nicolas Jacquemet, Béatrice Boulu-Reshef, Daniele Caliari, Philippe Colo, Quentin Couanau, Guillaume Pommey, Julien Combe, Shaden Shebayek, Rémi Yin, Justine Jouxtel, and Antoine Hémon. I want to thank seminar participants of the Roy seminar, TOM seminar, and Workgroup on experiments at Paris School of Economics. I would also like to thank seminar participants at Paris Nanterre University as well as conference participants and discussants at AFSE 2018, BRIC 2018, ASFEE 2018, Queen Mary University Ph.D. Workshop, FUR 2018, SARP 2018, ESA 2018 and Econometric Society European Meeting 2018. I acknowledge the support of ANR projects CHOp (ANR-17-CE26-0003) and DynaMITE (ANR-13-BSH1-0010) and Labex OSE (10-LABX-0093). All errors are mine.

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 in most real-life settings, choices are elicited using a *choice function*: the decision maker chooses one alternative from the choice set, i.e., the set of available alternatives. In principle, however, choices are commonly modeled with a *choice correspondence*: the decision maker chooses a non-empty set from the choice set.[1] Arrow (1959) gave the condition under which a choice correspondence, or a choice function, can be rationalized by a reflexive, complete, and transitive preference, i.e., a *classical preference*, the Weak Axiom of Revealed Preferences – WARP hereafter. Roughly, WARP states that if an alternative x is revealed better – i.e., preferred – to another alternative y, it cannot be the case that from another choice, y is revealed preferred to x.

In this paper, I introduce *pay-for-certainty*, a direct method for identifying choice correspondences. I allow decision makers to choose several alternatives, provide a small incentive for each alternative chosen, and then randomly select one for payment. Previous methodologies used a proxy to do so, preference for flexibility in Danan and Ziegelmeyer (2006) and Costa-Gomes et al. (2016), repeated choices in Agranov and Ortoleva (2017), or delegation to a random device in Qiu and Ong (2017), Sautua (2017), and Cettolin (2016). These methodologies might fail to identify indifference between two alternatives precisely.

Identifying a choice correspondence for each decision maker allows me to explore new questions. First, classical preferences often fail in practice to rationalize observed behavior. It is natural to look for which assumption might explain its failure. I study relaxations of completeness, transitivity of the indifference and menu-independence, using testable conditions on choice correspondences provided by Eliaz and Ok (2006) and Aleskerov, Bouyssou, and Monjardet (2007). Second, having identified choice correspondences, I can study indifference directly and assess how widespread it is. Third, with choice functions, decision makers have at most one maximal alternatives, whereas with choice correspondences, they might have more than one. The more extensive set of Pareto-superior alternatives opens up more room for agreement in collective decision making. Agreement between decision makers should be more accessible, as each has potentially several Pareto-superior alternatives.

Choice correspondences cannot be obtained directly. A 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 or incompleteness, as a simple example illustrates. Julia is a decision maker who likes equally coffee and tea. According to

[1] One 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.

this simple identification method, she can choose {coffee}, {tea} or {coffee, tea} in the choice set {coffee, tea}. Each chosen set has the same payoff for her. This creates a uniqueness problem. This procedure does not guarantee that the decision maker chooses all maximal – i.e., Pareto-superior – alternatives. On the other hand, revealed preference models (see, for instance, Sen (1971), Nehring (1997)) assume that the chosen set is the set of maximal alternatives.

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, Julia is better off by choosing {coffee, tea} and getting coffee or tea and 2ε in additional payment rather than choosing and getting {coffee} or {tea} and ε. The full characterization of *pay-for-certainty* is thus:

> In each set of alternatives, the decision maker chooses all the alternatives she wishes. She earns an additional payment of \$$\varepsilon$ by chosen alternatives. The alternative she gets is selected from her chosen alternatives using, for instance, a uniform random draw.

The additional payment of pay-for-certainty implies forgone gains when the chosen set is not the whole set. Choosing only one alternative earns ε, whereas choosing two earns 2ε, and so on. I show that under mild monotonicity conditions, the decision maker chooses all maximal alternatives. 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 ε of each other. The method shares many features with the experiment of Costa-Gomes et al. (2016), but there are some key differences. 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.

Under the assumptions of transitive strict preferences, monotonicity for the gains, and constrained maximization of the additional payment, pay-for-certainty (partially) identifies the choice correspondence. Importantly, I do not assume completeness of the preference nor transitivity of the indifference relation. Constrained maximization is not directly testable. Instead, I provide a testable implication of all the assumptions I make: *increasing chosen set*. Increasing chosen set captures the intuitive idea that, for a given choice set, the larger the additional payment is, the larger the chosen set should be. Hence, for a given choice set, the chosen set with pay-for-certainty with no additional payment is a subset of the set of maximal alternatives, while it is a superset of the set of maximal alternatives when the additional payment is positive. When increasing chosen set is verified, I *partially* identify the choice correspondence of the decision maker. I precisely identify maximal alternatives in a given set when the two are equal. I *fully* identify the choice correspondence of a decision maker when for all choice sets, I precisely identify all maximal alternatives. I will call the choice correspondence of the decision maker the fully identified ones, otherwise mentioning that the identification was only partial.

I have carried out an experiment to illustrate pay-for-certainty in the laboratory. The 223 subjects of the experiment chose between four tasks they had to perform at the end of the session. They were rewarded depending on their performance in the tasks. Subjects chose in all of the 11 possible subsets of tasks in order to fully identify their choice correspondence. Subjects chose three times in the same sets to test the two identification assumptions: twice with different strictly positive additional payments and once with no additional payment. The computer selected one of the 33 choices randomly for the payment and drew the task they performed from the chosen set randomly. I also ran an experiment with the same tasks but forcing the choice of singletons, which amounts to identifying a choice function. Finally, I varied the information provided about the tasks to study the influence of the quality and quantity of information on the choices.

In the experiment, subjects did not choose singletons when not constrained to do so. 60% of all choices were *not* singletons, and only 3% of subjects always chose singletons. Importantly, even without incentive to choose multiple alternatives, in the 0-correspondence case, the choice of non-singletons does not disappear. It also remained when the information provided was quite complete. Overall, I identify the choice correspondence of 26% of subjects and at least partially identify it for 72% of subjects.

Descriptively, reflexive, transitive and complete preferences together often fail to explain observed choices. Normatively, the appeal of some assumptions is also dubious. Aumann (1962) and Bewley (2002) have criticized complete preferences. Similarly, Armstrong (1939) and Luce (1956) criticized transitive indifference and introduced just-noticeable difference modeling in response. Sen (1997) was critical of menu-independent (i.e., set-independent) maximization. In the last decade, the theoretical literature on incomplete preferences blossomed, following Ok (2002) and Bewley (2002), but few empirical studies have been conducted.[2] Similarly, just-noticeable difference and menu-dependent model have seldom been explored.[3] Identifying the choice correspondence of a decision maker allows me to explore these assumptions.

First, classical preferences rationalize 98% of (fully identified) choice correspondences, leaving almost no room for alternative modeling of preferences. It represents only 26% of the subjects of my experiment, however. For the remaining ones, I will explore the alternative models on their choices with additional payments, which I will call hereafter the ε-correspondence. Only 46% of the ε-correspondences satisfy WARP. I use choice correspondences to study incomplete preferences directly, using the Weak Axiom of Revealed Non-Inferiority – WARNI hereafter. It is equivalent to rationalizability by a reflexive and transitive preference, a corollary to Eliaz and Ok (2006)'s Theorem 2. Incomplete preferences rationalize 49% of ε-correspondences. Second, I study just-

[2] To the best of my knowledge, five experiments are looking at incomplete preferences (Danan and Ziegelmeyer (2006), Cettolin (2016), Costa-Gomes et al. (2016), Qiu and Ong (2017), Sautua (2017)). Only one of them, by Costa-Gomes et al. (2016), looked at alternatives in the certain domain.

[3] Models of status-quo bias, attraction and decoy effects, and so on, are menu-dependent models of choice, but they always assume the impact of the menu-effect. Here I am thinking of the effect of the menu *per se*, not assuming the mechanism behind the effect of the menu.

noticeable difference models, formalized by Luce (1956) and Fishburn (1970) among others. The "coffee and sugar" example[4] captures the intuition behind just-noticeable difference models. The difference between the two alternatives might be imperceptible, but the addition of imperceptible differences might be perceptible. I use Aleskerov, Bouyssou, and Monjardet (2007) testable implications of just-noticeable difference. In my experiment, just-noticeable difference and incomplete preferences rationalize the same choice correspondences and ε-correspondences. Both rationalize only marginally more than classical preferences. This result is close to the proportion obtained by Costa-Gomes et al. (2016), in the context of sure outcomes, but in contrast with the rest of the literature (Danan and Ziegelmeyer (2006), Cettolin (2016), Sautua (2017) and Qiu and Ong (2017)) which use risky or ambiguous lotteries to study incomplete preferences.

Finally, I relax menu-independence. I study the models of Frick (2016) and Aleskerov, Bouyssou, and Monjardet (2007). They are variations on just-noticeable difference models, but now the threshold to be noticed might depend on the menu and the alternatives considered. The easiest one to satisfy, the context-dependent model of Aleskerov, Bouyssou, and Monjardet (2007), rationalizes up to 93% of ε-correspondences, while not being a void requirement on choice correspondences.[5] It reveals reflexive, complete and transitive preferences. In this model, however, decision makers might sometimes choose sub-optimal alternatives.

Identifying the choice correspondence also allows me to study indifference. It matters as a decision maker might violate WARP with choice functions if she has classical preferences with indifference. The following example illustrates why.

Example: Singleton choice Julia is indifferent between three hot drinks, coffee, tea, and hot chocolate, and she randomizes when she is indifferent and forced to choose one alternative. I identify her choice function, by forcing her to choose singletons. Her observed choices are: c({coffee, tea, hot chocolate}) = {tea}, c({coffee, tea}) = {coffee}, c({coffee, hot chocolate}) = {hot chocolate}, and c({tea, hot chocolate}) = {tea}. This choice pattern does not satisfy WARP, and the revealed preferences of Julia are coffee$\sim$tea, tea$\sim$hot chocolate, and coffee$\succ$hot chocolate.[6] Singleton choices do not reveal Julia's preferences, and I deem her "irrational" because I have forced her to choose singletons.

Forcing decision makers to choose one alternative can generate this kind of error. Danan (2008) explains under which conditions on the behavior of the decision maker a choice function successfully assesses her rationality.

[4]Luce (1956): "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, $i=0,1,\ldots,400$, 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 $i=400$."

[5]See Appendix C.3.1 for a study of the explanatory power of the different models on choice correspondences.

[6]There are different possibilities to assess revealed preferences, as shown in Sen (1971). I use Arrow (1959)'s definition of revealed preferences.

Experimentally, the restriction of choice to singletons is not what explains violations of WARP, as 46% of ε-correspondences and 57% of all singleton choices satisfy WARP. Choices with ε-correspondences are more difficult than with singletons, which probably explained the insignificant decrease in the satisfaction of WARP.[7] On the other hand, the restriction to singleton choice significantly changes the revealed preferences. On average, for ε-correspondences and choice correspondences that satisfy WARP, 50% of comparisons between alternatives are indifference relations. If I use preferences obtained with the context-dependent model of Aleskerov, Bouyssou, and Monjardet (2007), when it rationalizes ε-correspondences, 36% of comparisons are indifference relations. Indifference relations have a significant share of revealed preference relations. It is likely to imply that the number of Pareto-superior alternatives for each decision maker is more substantial than usually assumed, which is relevant for individual welfare analysis.

This identification of indifference has positive consequences for collective welfare analysis as well. By allowing subjects to have more than one maximal alternatives, choice correspondences facilitate the agreement among decision makers compared to choice functions. 73% of choice correspondence who satisfy WARP agree that the addition task and the spellcheck task are the best alternatives. This proportion is much higher than when I identify choice functions which satisfy WARP. In that case, the most preferred agreed upon task is the copy task, and 31% of subjects prefer it. These results show that in practice, an agreement between decision makers is more accessible with choice correspondences. It is a result in the spirit of the one proved by Danan, Gajdos, and Tallon (2013) for incomplete preferences.

The paper is structured as follows. Section 1 characterizes the domain of application of the pay-for-certainty method, as well as testable consequences of the assumptions I make. Section 2 describes the experiment. Section 3 studies the identification of choice correspondences in practice. Section 4 studies classical preferences and then alternative rationalization of choice correspondences. Section 5 explores the prevalence of indifference. Section 6 explores the aggregation of preferences and choices with choice correspondences. Section 7 concludes.

1 Method

In this section, I describe the *pay-for-certainty* method and lay down the conditions under which it allows me to identify the choice correspondence of a decision maker.

[7] If we compare all possible choice correspondences and all possible choices functions and pick one at random in each set, a choice function is 400 times more likely to satisfy WARP as shown in Appendix C.3.1.

1.1 Setup

Formally, I start with X, a *finite* set of alternatives.[8] I model choices made by the decision maker using a correspondence on X, that is, a mapping that associates to any subset of X another subset of X. Each non-empty subset S of X is a *choice set*. A *choice correspondence* c on X associates to each choice set a non-empty subset of that set.

$$\begin{aligned} c: 2^X \backslash \emptyset &\rightarrow 2^X \backslash \emptyset \\ S &\rightarrow S' \subseteq S \end{aligned}$$

Alternatives in $c(S)$ are the ones *chosen* in S, whereas alternatives in $S \backslash c(S)$ are *unchosen* in S. For incentive purposes, the difference between chosen and unchosen alternatives is that the decision maker never gets unchosen alternatives whereas she gets *one* of the chosen alternatives. In this sense, the decision maker wants alternatives from $c(S)$ and does not want alternatives from $S \backslash c(S)$.

A *choice function* is a particular choice correspondence: it associates to each choice set a *unique* alternative from the choice set.

$$\begin{aligned} c: 2^X \backslash \emptyset &\rightarrow X \\ S &\rightarrow x \in S \end{aligned}$$

Choices *per se* are not very useful for out-of-sample predictions or individual welfare analysis. A rationale for the choice is needed to do so. One widely used hypothesis in economics is to assume that decision makers behave *as if* they have a preference over alternatives, a preference that drives their choice. Revealed preferences construct preferences from choices. In theory, with choice correspondences, the assumption is that all chosen alternatives are the best in the choice set (see Sen (1971) for instance). That is, chosen alternatives are strictly better than unchosen ones. An exception is Eliaz and Ok (2006), who assume that chosen alternatives are not worse than unchosen ones, and potentially not comparable. With choice functions, the assumption is usually that the chosen alternative is *not worse* than unchosen alternatives. That is, chosen alternatives are weakly better than unchosen ones. The first question I will tackle is to understand under which conditions the assumption of strict preferences on choice correspondences holds empirically.

Assumption 1.1. *The decision maker behaves as if she has* reflexive *preferences* $\succeq$ *on* X*. The strict part of the preference,* $\succ$*, is* transitive*.*

In other words, $\succeq$ is a preorder. I do not assume that preferences are *complete*, only that they are transitive and reflexive. Importantly, I do not assume the indifference part to be transitive. It

[8] The design does not extend readily to infinite sets. I discuss why in Appendix A.2.

allows me to explore models of intransitive indifference, i.e., just-noticeable-difference, as in Luce (1956) and Fishburn (1970).

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. I will denote $M(S)$ the set of maximal alternatives to discuss these assumptions.

Definition 1.1 (Set of Maximal Alternatives in S ($M(S)$))**.**

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

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 and therefore is a strict partial order, which is acyclic. $M(S)$ 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 $S \backslash M(S) = D(S)$ the set of *dominated* alternatives, as she could choose a better alternative. That is, the set of chosen alternatives is a subset of the set of maximal alternatives: $c(S) \subseteq M(S)$.

Most theoretical work (see Aleskerov, Bouyssou, and Monjardet (2007) p29-30; Sen (1997); Nehring (1997) among others) adopt the following stronger statement. The set of chosen alternatives *is* the set of maximal alternatives: $c(S) = M(S)$. This stronger interpretation 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 $c(S)$:

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

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

While very convenient, this 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 for the method I am introducing now.

1.2 Pay-for-Certainty

The objective of the method I am introducing, pay-for-certainty, is to recover the set of maximal alternatives of a decision maker in a set S. In experiments, decision makers usually have to choose according to a choice function. For each choice set, the decision maker chooses precisely one alternative, which is then given to her. Arguably, this is close to the situation in the field, where a decision maker generally chooses 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) has shown that forcing choice modifies the choice of decision makers. Danan and Ziegelmeyer (2006) and Costa-Gomes et al. (2016) 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. This evidence implies that decision makers like some flexibility in their choices. Choice correspondences introduce this flexibility in experiments, by not forcing decision makers to select precisely one alternative.

With choice correspondences, for incentive purposes, precisely one of the selected alternative must be used for the payment. I select this alternative through a *selection mechanism*. It associates to a set of alternatives one alternative from this set. I discuss various selection mechanisms and their interactions with the possibility to choose a set in Appendix A.3.1. For now, consider the *uniform selection mechanism*:

> From the set of chosen alternatives, the alternative the decision maker gets at the end is selected 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.

Contrary to choice functions elicitation, there are different ways to elicit a choice correspondence. The simplest is what I call the *naive choice correspondence elicitation*:

> In a choice set, the decision maker chooses a non-empty subset. A selection mechanism selects the alternative.

The naive elicitation procedure has one problem. It does not guarantee that decision makers choose all maximal alternatives. This non-uniqueness, non-maximal problem is very general and arises for a classical decision maker because of indifference. The naive choice correspondence elicitation procedure can only guarantee that $c(S) \subseteq M(S)$. One solution to elicit indifference dates back to Savage (1954). Danan (2008) formalizes it: costly strict preferences. The intuition is straightforward: 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 she is indifferent between alternatives. I built the *pay-for-certainty elicitation* on this intuition.

> From the choice set S, each alternative chosen adds an *additional payment* of $\varepsilon > 0$ per alternative to the gain of the decision maker.[9] The total additional payments are $|c(S)| \times \varepsilon$. The alternative she gets is selected using a selection mechanism.

The selection mechanism is not *a priori* specified, but in practice, I 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 her preference is so weak that the difference is hardly perceptible, she will choose $\{x, y\}$ and I will think she is indifferent between x and y, which is not the case. If ε is large enough, choosing $\{x, y\}$ and getting 2ε is better than choosing $\{x\}$ (or $\{y\}$) and getting ε. 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 by ε.

In the next sections, I will study *under which condition the pay for certainty method allows one to recover the set of maximal alternatives.* I make two kinds of assumptions to do so: assumptions on preferences and assumptions on behavior.

1.3 Assumptions on Preferences

The main feature of pay-for-certainty compared to the naive choice correspondence elicitation is the additional payment for each alternative chosen. Observed choices are on the alternatives augmented by the payment, not on the original set of alternatives. The preference I want to elicit however is on the original set of alternatives. I have to extend the set-up to take into account the payment and link it with the original set of alternatives.

I extend the setup to $X \times \mathbb{R}$ to account for the payment. An element in this set is a couple (x, r). If r is positive, it is interpreted as a payment *to* the decision maker, if it is negative, as a payment *from* the decision maker. Denote $\succeq_2$ the preferences of the decision maker on this new set of alternatives. For clarity, I call preferences on X *unobserved* preferences, whereas I call preferences on $X \times \mathbb{R}$ *observed* preferences.[10]

I impose some structure on preferences on $X \times \mathbb{R}$ with two assumptions, *monotonicity* and *transitivity*, and link preferences on X and preferences on $X \times \mathbb{R}$ with one assumption, *identity*.

[9]The 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.

[10]Observed and unobserved are for the time being a label. At the end of the method part, the label will be more transparent.

Assumption 1.2 (Identity)**.** *Observed and unobserved preferences are the same when payments associated with the alternatives are null.*

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

Assumption 1.3 (Monotonicity)**.** *If the only difference between two alternatives in $X \times \mathbb{R}$ is the payment, the decision maker always prefers the highest payment to the lowest one.*

$$\textit{for all } x \in X, \textit{for all } r, r' \in \mathbb{R}, r > r' \Leftrightarrow (x, r) \succ_2 (x, r')$$

Assumption 1.4 (Transitivity of observed strict preferences)**.**

$$\textit{for all } x, y, z \in X, \textit{for all } r, r', r'' \in \mathbb{R}, (x, r) \succ_2 (y, r') \textit{ and } (y, r') \succ_2 (z, r'') \Rightarrow (x, r) \succ_2 (z, r'')$$

Using identity, transitivity of observed strict preferences implies transitivity of unobserved strict preferences, which is Assumption 1.1. The converse is true only without payment. When a sufficiently large payment is added to one of the incomparable alternatives, the payment will drive the comparison, and the alternatives will become comparable. In the paper, I consider only small payments so that incomparable alternatives according to unobserved preferences should remain incomparable alternatives even with an additional payment.

1.4 Assumptions on Behavior

This section lays out the assumptions on behavior needed to elicit the choice correspondence.

Definition 1.2 (Strict r-domination)**.** An alternative x r-dominates another alternative y if and only if $(x, 0) \succ_2 (y, r)$.

A decision maker chooses one alternative over the other with the pay-for-certainty procedure when the strict preference between the two is large enough. r-domination precisely defines this difference. The decision maker prefers enough x to y so that she still prefers x without payment to y with a payment r.

Remark. Assuming Monotonicity, Transitivity, and $r > r'$, if x r-dominates y, it also r'-dominates y.

The counterpart to the notion of maximal alternatives in X is r-maximal alternatives in $X \times \mathbb{R}$.

Definition 1.3 (r-maximal alternatives in S)**.** $M_r(S)$ is the set of r-maximal alternatives in S which are all the alternatives that are not strictly $\frac{1}{|S|}r$-dominated.

$$M_r(S) = \left\{ x \in S \middle| \text{there is no } y \in S, \left(y, \frac{1}{|S|}r\right) \succ_2 (x, 0) \right\}$$

M_r is the set of sufficiently preferred alternatives. Even adding a payment to another alternative in S is not enough to make it better than alternatives in $M_r(S)$. I restrict the payment associated with alternatives. Here, I consider that all alternatives have the same associated additional payment $\frac{1}{|S|}r$. The reason for these more specific payments will be more apparent later.

When the payment is null, by identity, $M(S) = M_0(S)$. For any positive payment r, $M(S) = M_0(S) \subseteq M_r(S)$. Consider any negative payments r. When the decision maker is indifferent between two alternatives x and y, she has an incentive to choose only one of x or y, to avoid incurring the cost. So for any negative r, $M_r(S) \subseteq M_0(S)$. Therefore, when an observed preference satisfies transitivity and monotonicity, I have for all $r > r'$, $M_{r'}(S) \subseteq M_r(S)$.

I have so far been quite general in the definition of payments associated with alternatives. The following is more restrictive, as I get closer to implementing the elicitation procedure in experiments. I start by defining the choice correspondence on the set of alternatives with payments. In doing so and for simplicity, I will constrain the payments and the relations between them.

Definition 1.4. (ε-correspondence) The *ε-correspondence* c_ε is the choice correspondence obtained on $2^X \backslash \emptyset$ when the payment for choosing an alternative in S is equal to $\frac{1}{|S|}\varepsilon$.

I have linked above $M_\varepsilon(S)$ and $M(S)$. I will now do something similar between c and c_ε, under the following assumptions.

Assumption 1.5 (Choice of non dominated alternatives)**.** *In each choice set S, for any positive payment, the decision maker only chooses non $\frac{1}{|S|}\varepsilon$-dominated alternatives.*

$$\text{for all } \varepsilon > 0, c_\varepsilon(S) \subseteq M_\varepsilon(S)$$

This mild assumption is weaker than the traditional equality assumed in the literature. Decision makers do not choose alternatives they view as worse than other available ones. Stopping at this assumption means that I use weak revealed preferences, which does not restrict observations much, as shown in Appendix C.3.1. I want to use the stronger version of revealed preferences, as is usual in the theoretical literature.

Assumption 1.6 (Constrained maximization of the payments)**.** *The decision maker maximizes the additional payment in the set of ε-maximal alternatives.*

$$c_\varepsilon(S) = \max_{S' \subseteq M_\varepsilon(S)} \{|S'| \times \varepsilon\}$$

An immediate consequence of Assumption 1.6 is that $c_\varepsilon(S) = M_\varepsilon(S)$ for all S. The introduction of the payment for choosing alternatives introduces a new incentive for decision makers. Importantly, an observer knows the incentive: decision makers should choose larger sets. She chooses non-maximal alternatives only up to the point where the payment compensates their introduction. These assumptions lead to the widely used interpretation of choice correspondences when the additional payment is strictly positive. Constrained maximization of the payments has no bite when the payment is null. In other words, I still have that $c_0(S) \subseteq M_0(S) = M(S) = c(S)$, while I cannot guarantee that $c_0(S) = M(S) = c(S)$.

I provide now a testable property for at least partially identifying the choice correspondence. This new – mild – consistency requirement is called *increasing chosen set.*

Property 1.1 (Increasing Chosen Set (ICS))**.** *A family of choice correspondences* $(c_\varepsilon)_{\varepsilon\in\mathbb{R}}$ *satisfies* increasing chosen set *if*

$$\textit{for all } S \in 2^X, \textit{for all } \varepsilon, \varepsilon' \in \mathbb{R}, \varepsilon < \varepsilon' \Rightarrow c_\varepsilon(S) \subseteq c_{\varepsilon'}(S)$$

Two ε-correspondences c_ε and $c_{\varepsilon'}$, with $\varepsilon < \varepsilon'$, follow increasing chosen set *for a given set S* if the chosen set in S with a payment ε is a subset of the chosen set in S with a larger payment ε'. Two ε-correspondences c_ε and $c_{\varepsilon'}$, with $\varepsilon < \varepsilon'$, follow increasing chosen set if they follow increasing chosen for all sets $S \in 2^X \backslash \emptyset$.

Increasing chosen set captures the idea that the higher the payment for choosing an alternative is, the larger the chosen set of the decision maker should be. Increasing chosen set is a consistency requirement across ε-correspondences with different payments, a difference with traditional consistency requirements of revealed preferences, such as WARP. The latter is a requirement within an ε-correspondence, not across different ε-correspondences.

Increasing chosen set is observable by varying the payments, contrary to the assumptions on preferences and behavior. Besides, increasing chosen set is a consequence of several of the assumptions above.

Proposition 1.1. *If a decision maker follows constrained maximization of the payments, monotonicity, and transitivity, her ε-correspondences follow increasing chosen set.*

The proof of this result is in Appendix A.1.

1.5 Identification of the Choice Correspondence

It is now time to sew together the pieces I have built so far. Under the constrained maximization of the payments, monotonicity, identity and transitivity assumptions, for any strictly positive payment ε:

$$\text{for all } S \in 2^X, c_0(S) \subseteq M_0(S) = M(S) = c(S) \subseteq M_\varepsilon(S) = c_\varepsilon(S)$$

These inclusions have two implications. First, when increasing chosen set is satisfied between a c_0 and c_ε with $\varepsilon > 0$, I can bound the choice correspondence c by subsets and supersets:

$$\text{for all } S, c_0(S) \subseteq c(S) \subseteq c_\varepsilon(S) \tag{1}$$

When the ε-correspondences satisfy increasing chosen set, I *partially identify* the choice correspondence of a decision maker.

Second, in this situation, I might have some sets S for which I have $c_0(S) = c_\varepsilon(S)$. For that set, $c(S) = c_0(S)$. I *fully identify* the choice correspondence when choices satisfy Equation (1) for all S with *equalities*. With a finite set of alternatives X, full identification happens when partial identification happens, and ε is small enough.

2 Experiment

In the previous section, I have provided one testable property to assess the validity of pay-for-certainty for the choice correspondence elicitation, increasing chosen set. I now illustrate pay-for-certainty and the validity of this condition by running an experiment in the laboratory. This section introduces the experimental design.

2.1 Description of the experiment

In order to identify a choice correspondence, I need at least two ε-correspondences, one without payment and another with a strictly positive payment. I want to compare empirically choice correspondences and choice functions. I also elicit choice functions in the experiment to do so. I feared some priming between the elicitation of ε-correspondences and choice functions. I have thus separated both elicitation, and the differences are between subjects.[11]

I need choices from all possible subsets of the grand set of alternatives to be able to falsify the different models I will consider. I am 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, I have to study 11 choices, for five, 26 choices, and for six, 57. In practice, X could contain 4 or 5 alternatives. I wanted the subjects to repeat the same choices for at least two different payments, and thus I settled on a set of 4 alternatives.

Finally, I wanted to study incomplete preferences. Two possible drivers of incomplete preferences are the lack of information about the alternatives and the lack of information about the subjects' preferences.[12] Assessing the knowledge of subjects about their preference is hard. Instead, I settled on varying the information they were provided about the alternatives, an idea close to what Costa-Gomes et al. (2016) did.

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 experiment lasted between 40 and 60 minutes, depending on the different treatments and the speed of the subjects.

2.1.1 Choice objects

Subjects chose between four different paid tasks:

- 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.[13] They earned 10 cents for each mistake corrected and lost 10 cents for each mistake added.
- 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.
- 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.

[11]Indeed, I specifically feared that because the choice function made subjects think of one alternative in each set, an ε-correspondence elicited after would exhibit smaller chosen sets.

[12]Another 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.

[13]For 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 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. I selected the task they performed by drawing one of the 33 choices they made at random. From this chosen set, I uniformly drew one task.

Additions and sequences were randomly generated and thus did not have an end. I told subjects that it was not possible to finish the spell-checking task in less than three minutes. 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 treatment, always done after they had made all their choices.

2.1.2 Timing

Subjects in the experiment went through five steps, which I detail in the next subsections. First, I read the instructions about the experiment (available in Appendix B.1) to the subjects. This part included a description of the tasks they had to choose. This description differed across the treatments. Each subject also had a printed version of the instructions available during the whole session. Second, subjects chose three times eleven choices according to pay-for-certainty, at different payment levels. For some, I replaced the first payment by choices of single alternatives. Third, I measured 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.

2.1.3 Treatments

To investigate the influence of information provided on the size of the chosen sets, I varied the explanations of the tasks. I 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 what I have given above. In the *video* treatment, subjects first received the sentence treatment. Then I showed them a video explaining each task. The video showed the interface of the task and explained how to perform it. Another one replaced the text used for spell-checking in the real task. The choice function elicitation followed this treatment. Finally, in the *training* treatment, subjects first went through the video treatment. Then they trained on the tasks for 30 seconds. 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. In addition to the quantity of information driving the size of the chosen sets, I expected it to influence incompleteness. I took the video treatment as the baseline treatment, the other two showcasing the influence of information on choices.

2.1.4 Choices

I 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. I studied three different additional payment levels:

- no gain: 0 cents (which is the naive elicitation described earlier);
- low gain: 1 cent;
- high gain: 12 cents.

Following pay-for-certainty, in each set, choosing an alternative implied a gain of $\frac{1}{n}\varepsilon$ where n is the size of the choice set. For instance, if the choice set is of size 2 and the subjects faced a high gain, choosing an alternative pays 6 cents. Fractions of cent where paid by randomization: 0.25 cents corresponds to a 25% probability of getting 1 cent and a 75% probability of getting 0 cents. When the subjects chose several alternatives in a set, the computer used the uniform selection mechanism to select the alternative eventually given to the subjects. The 11 choices of a given level of payment were performed in a row to avoid confusion between the different additional payment levels. The order of the different payment level was random. The order of the different choice sets was random. The use of a proportional rather than a linear payment scheme for adding chosen alternatives is a consequence of the externality of adding an alternative. In addition to increasing the gains, it decreases the probability of other chosen alternatives to be drawn at the end. This probability variation depends on the size of the choice set.

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. I did that to decrease the risk of errors from choosing too quickly. Screenshots of a choice screen followed by a confirmation screen are in Appendix B.2. Subjects were reminded of the selection mechanism on each screen.

There are two differences for subjects who chose according to a choice function. Their first 11 choices were the choice function. Subjects made the 22 next choices according to pay-for-certainty with no and low gains, the order of the gain was random.

2.1.5 Measure of Risk-Aversion

I elicited risk-aversion as well. I used Dohmen et al. (2011) method to assess the certainty equivalent of the lottery (5€, 1/2; 0€, 1/2), by increments of 20 cents between 0€ and 3€. The average certainty equivalent was 1.98€. 78% of subjects had consistent certainty equivalent measure: they had precisely one switching point between the lottery and the certain amounts.[14] Their certainty equivalent is 1.94€, not significantly different from the 2.14€ of subjects who do not have consistent risk-aversion.

Risk-aversion does not play a role in my experiment. There is no significant difference in the results when we control for risk aversion or stratify by risk aversion levels. I show some of these robustness checks in Appendix B.4, as well as the distribution of certainty equivalents.

2.1.6 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.

In addition to this information, I asked subjects which tasks they liked or disliked. I ordered the tasks accordingly: preferred ones, worst ones and "in between" ones. When the subjects provided no information, all alternatives were considered "in between". I built the corresponding choice correspondence. In Appendix C.1, I study the relationship between stated preferences and revealed preferences.

2.2 Subjects

The sessions took place between the 14th of November 2017 and the 29th of May 2018. The earliest started around 11 am 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.[15]

51 subjects participated in the choice function elicitation, during three separate sessions. After the choice function elicitation, they all chose according to pay-for-certainty with no and low gains and video treatment. For the first two sessions, significant priming happened, that I was able to correct by explaining more carefully the choice process.[16]

119 subjects participated in the pay-for-certainty elicitation procedure with no and low gains and video treatment. 102 did so with high gains. 33 subjects participated in the pay-for-certainty elicitation procedure with all gains and the sentence treatment. 37 subjects participated in the pay-for-certainty elicitation procedure with all gains and the training treatment.

[14] For subjects with more than one switching point, I used the latest switching point to compute the certainty equivalent, as is standard in the literature on risk aversion.

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

[16] I thus dropped this data.

The 223 subjects of the sample have done the measure of risk aversion, the questionnaire, and the tasks.

The demographics of the sample, shown in Appendix B.3 shows that it is neither a representative sample of the population nor a typical student pool.

3 Identification of the Choice Correspondence

I study in this section the identification of choice correspondences. I also show the choices obtained with ε-correspondences and the difference with choice functions.

3.1 Non-Singleton Chosen Sets

A first natural question is whether subjects chose non-singletons in ε-correspondences. Table 1 shows the size of the chosen sets depending on the size of the choice sets. Overall ε-correspondences, 59.7% of chosen sets are not singletons. Forcing subjects to choose only one alternative is a restriction in these cases. Only 8.2% of all ε-correspondences are choice functions. That is, subjects chose singletons in all choice sets. 3% of the subjects always chose as if I have forced them to choose singletons, that is, all their ε-correspondences are always made of singletons. For these subjects, eliciting a choice function is not a restriction.

Table 1 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. I study menu-dependent – i.e., set-dependent – choice in Section 4.4.

Table 1: Proportion of chosen sets of different sizes for all ε-correspondences. Horizontally are the size of the chosen sets and vertically of the choice sets.

Set size	1	2	3	4
2	53.3%	46.7%		
3	26.6%	41.5%	31.9%	
4	16.7%	26.2%	34.9%	22.2%

The incentive to choose more than one alternative introduced by pay-for-certainty is one driver of non-singleton choice, but not the only one. Table 2 shows that even with no payments, subjects do not always choose singletons. The introduction of the additional payment has a non-linear effect on choices: a small gain significantly changes the aggregate proportion of singletons chosen, but small variations inside the gain domains do not. 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. I cannot disentangle the two explanations in the experiment. This aggregate behavior is consistent with increasing chosen set but is incompatible with the full identification of the choice correspondence for all subjects.

Table 2: Proportion of singletons chosen at different payments, depending on the size of the choice set. Differences between no gains and positive gains are significant. Differences between low and high gains are not.

Choice set size	No	Low	High
2	65.6%	47.3%	46.5%
3	39.9%	20.5%	18.6%
4	27.0%	11.1%	11.6%

Table 3 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.

Table 3: Proportion of chosen sets that are singletons, depending on the information and the size of the choice set.

Choice set size	Sentence	Video	Training
2	40.1%	54.8%	60.8%
3	15.4%	28.9%	29.5%
4	8.1%	18.8%	18.0%

The non-singleton 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.

3.2 Increasing Chosen Sets

I have shown in Section 1 that increasing chosen set is key to assess the validity of the pay-for-certainty method. When a subject satisfies increasing chosen sets, I can at least partially identify her choice correspondence. Increasing chosen set is a within subjects across ε-correspondences consistency requirement. I have at most three pairs of additional payments for each subject: no and low, no and high and low and high. To identify the choice correspondence, however, only the two pairs no and low payments and no and high payments matter. Because I am mainly interested in the identification of the choice correspondence, I only show in Table 4 results between no and low payments and between no and high payments.

Table 4 shows that I identify the choice correspondence for 26% of the subjects, and partially identify it for 72% of them. Comparing the different treatments, increasing chosen set is overall significantly more satisfied in the sentence treatment than in the other two. It is due to a higher prevalence of fully indifferent subjects, as shown later in Section 5. These subjects always choose the whole set and therefore trivially satisfy increasing chosen sets. The training treatment provides better identification of the choice correspondence, but the difference is not significant.

Table 4: Proportion of subjects following increasing chosen set for different pairs of additional payments and different information.

Treatments	(no, low)	(no, high)	partially	fully
Sentence	82%	83%	88%	27%
Video	55%	60%	68%	24%
Training	59%	57%	73%	32%
All	60%	63%	**72%**	**26%**

The 28% of subjects who do not satisfy increasing chosen set violates the consistency requirement of the experiment. I cannot even partially identify their choice correspondence. Assessing the reason behind these violations is difficult. Maybe subjects chose at random – this is very unlikely, as will be shown in Section 4, maybe they are tired of making broadly similar choices and make some mistakes, and so on. One way to take into account the latter is to allow for some mistakes explicitly. If I allow subjects to violate increasing chosen sets in one set, the satisfaction rate rises to 85%, and if I allow for two mistakes, to 93%.

These numbers suggest that overall subjects understood the pay-for-certainty procedure and are generally consistent with the assumptions laid out in Section 1. Increasing chosen set is a robust property, satisfied with different treatments and almost always satisfied when I take into account the possibility of one or two mistakes.

From now on, I give the results for three different samples:

- Choice correspondences, with 49 observations;
- Partially identified choice correspondences, with 137 observations. For simplicity and consistency, I use the choice obtained via pay-for-certainty with no payments for these. This choice is arbitrary; I could have presented the results with positive payments. The results are not significantly different. This sample also includes the sample of identified choice correspondences.
- All ε-correspondences, 550 observations.

4 Reflexive, Complete, Transitive and Menu-Independent Preferences, and Beyond

A decision maker behaves *as if* she has reflexive, transitive and complete preferences on X if and only if her choice correspondence satisfies the Weak Axiom of Revealed Preferences on 2^X, as shown by Arrow (1959). This result holds if and only if the choice correspondence is defined over the powerset, which is the case in my experiment. I borrow the formulation of WARP from Eliaz and Ok (2006).[17]

Axiom 4.1 (Weak Axiom of Revealed Preferences (WARP))**.** For any $S \in 2^X$ and $y \in S$, if there exists an $x \in c(S)$ such that $y \in c(T)$ for some $T \in 2^X$ with $x \in T$, then $y \in c(S)$.

In words, if x is available in T and y is chosen, one may conclude that y is at least as good as x for the decision maker, and thus whenever x is chosen from a set S that contains y, so must y. The as if preference maximized by the decision maker is obtained using revealed preferences.

- x is *weakly revealed preferred* to y (denoted xRy) if and only if there is a set in which x is chosen, and y is available.

$$xRy \Leftrightarrow \text{there exists } S \in 2^X, x \in c(S), y \in S$$

- x is *strictly revealed preferred* to y (denoted xPy) if and only if there is a set in which x is chosen, and y is not.

$$xPy \Leftrightarrow \text{there exists } S \in 2^X, x \in c(S), y \in S \backslash c(S)$$

[17] WARP has taken various names in the literature. It is Arrow (1959)'s Condition 4, Chernoff (1954)'s postulate 10 and Aleskerov, Bouyssou, and Monjardet (2007)'s Arrow's Choice Axiom.

- x is *revealed indifferent* to y (denoted xIy) if and only if there is a set in which both x and y are chosen.

$$xIy \Leftrightarrow \text{there exists } S \in 2^X, x \in c(S), y \in c(S)$$

R, P and I are binary relations, that is, subsets of X^2. When a decision maker satisfies WARP, P and I are a partition of R. In revealed preferences terminology, WARP states that it cannot be that both x is strictly revealed preferred to y and y is weakly revealed preferred to x, i.e., if xPy, then not yRx. It is easy to know which alternative is the best when WARP is satisfied, use the revealed preferences as the preferences of the decision maker.

4.1 WARP in Practice

Overall, 46% of ε-correspondences satisfy WARP and 57% of choice functions do.[18] The difference in satisfaction between choice functions and ε-correspondences is not significant. Only 19% of the subjects satisfy WARP for their 0, 1 and 12-correspondences together, whereas 25% never satisfy it. For subjects whose choice correspondence is partially or fully identified, the satisfaction of WARP rises to 58% and 98%. These numbers are lower than increasing chosen sets, indicating that some subjects are consistent across choice correspondences, but not within them.

Table 5 shows the satisfaction of WARP across the different treatments for ε-correspondences. The sentence treatment has the highest rate of satisfaction. As shown in Table 11 in Section 5, in the sentence treatment, subjects have more indifference relations. They choose more often the whole set, trivially satisfying WARP in the process. The information provided influences the quality of choice. Subjects earned more on average when they had more information, even though they satisfied WARP less.

Table 5: WARP and average gains by information, significance levels are assessed compared to the video treatment. For gains, the sentence and the training treatment are significantly different at the 1% level.

Treatment	Sentence	Video	Training
WARP	62%***	44%	45%
Gain in the Task	2.66**	3.23	3.88**

WARP and increasing chosen sets are two orthogonal consistency requirements, within and across ε-correspondences. There is no reason *a priori* for them to be correlated, but they happen to be. Each subject has two or three ε-correspondences. I can, therefore, assess for each subject whether

[18] One can wonder whether the different models I will present all along have explanatory power. The answer is yes, and I show in Appendix C.3.1 that none is trivially satisfied by a random choice correspondence or choice function. Incidentally, it proves that subjects have not chosen at random in my experiment.

some, all or none of their ε-correspondences satisfy WARP. Similarly, I can assess whether increasing chosen sets is satisfied for some, all or no additional payment pairs. Table 6 shows that these partial, full or absence of satisfaction of both axioms are correlated: 74% of the ε-correspondences are on the diagonal. Some subjects are more consistent than others, overall.

Few papers proceed to similar tests of WARP on experimental data. The setups can be largely different, but their results are not. They all show significant violations of WARP. Except Costa-Gomes et al. (2016), these paper all use choice functions to elicit preferences.

Table 6: WARP and increasing chosen set.

	All correspondences	Some correspondences	Never
All payments	**20%**	19%	2%
Some payments	1%	**36%**	7%
Never	0%	5%	**18%**

Bouacida and Martin (2017) have a section on experimental data and the satisfaction of WARP in an experiment that studies impatience. 47% of subjects satisfy WARP all the time, which is higher than the comparable figure of 19% found here, but they had only two choice functions per subjects, whereas here I have three choice correspondences, a more complex setup. On all ε-correspondences, the average rate of satisfaction of WARP is of 47% here and 64% in their experiment. 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. My subject pool is in between their two pools, as shown in Appendix B.3. Costa-Gomes et al. (2016) found that 28% of observed choice correspondences satisfy WARP, a figure lower than here.[19]

As I did with increasing chosen sets in Section 3.2, I can assess the severity of the violations of WARP. Figure 1 shows the minimal number of choice sets to remove in order to satisfy WARP, for different additional payments and the choice function.[20] The number of sets to remove is different between the choice functions and the ε-correspondences for the different payments at the 5% level, except for the high payment. As in the study of increasing chosen set, allowing for one or two mistakes, that is, removing one or two sets, increases the satisfaction of the axioms substantially. Indeed, 63% and 76% of ε-correspondences satisfy WARP with this relaxation. For choice functions, these numbers are respectively 82% and 88%.

[19] I give more details on this experiment in Section 4.2.1.

[20] I averaged over all minimal combinations when different combinations of sets to remove are possible to rationalize observed choices.

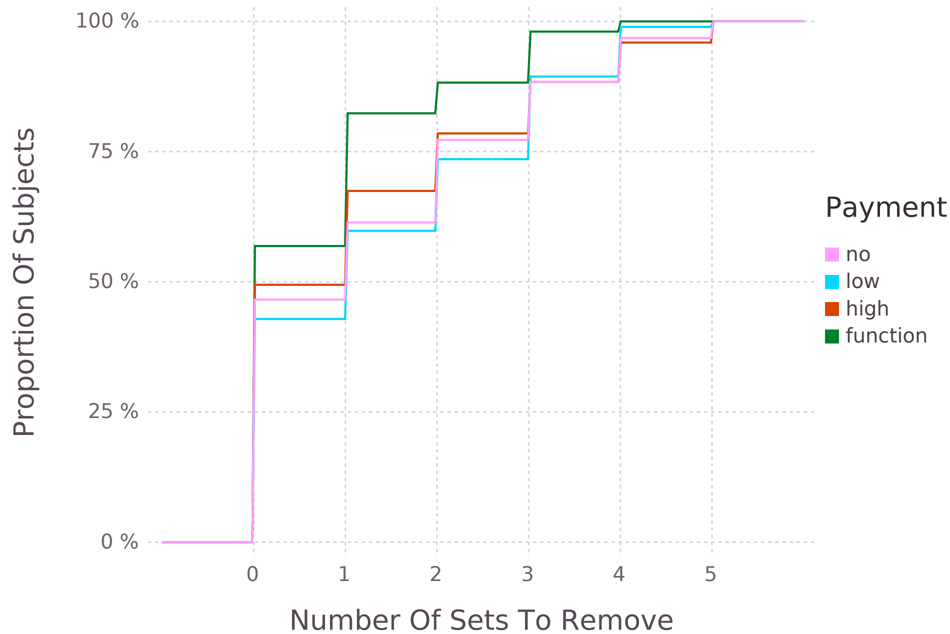

Figure 1: CDF of the number of sets to remove in order to satisfy WARP.

4.2 Incomplete and Intransitive Indifference Preferences

It is possible to rationalize observed choice correspondences when WARP fails with relaxations of complete and transitive preferences.

First, I relax completeness, leading to partial orders (PO-rationalizability). The critique of the normative appeal of the completeness of preferences assumption dates back to Aumann (1962). Eliaz and Ok (2006) provide the first criterion for rationalizability of observed choices by incomplete preferences. Aleskerov, Bouyssou, and Monjardet (2007) provide an equivalent axiomatization.

Second, I relax transitivity of the indifference part. Intransitive indifference dates back at least to Armstrong (1939). These models are known as models of intransitive indifference or just-noticeable difference. I study two of the most common ones, semi orders (SO-rationalizability) from Luce (1956) and interval orders (IO-rationalizability) from Fishburn (1970).

I can investigate these models because I have identified choice correspondences. On choice functions, they rationalize the same observed choices, as explained in Appendix C.3.1. To the best of my knowledge, I am the first to study their empirical validity. There are other investigations of incomplete preferences in economics, but they rely on the identification of incompleteness with another phenomenon. For instance, Danan and Ziegelmeyer (2006) identify incomplete preferences with a preference for flexibility.

4.2.1 Incomplete Preferences

In addition to being normatively doubtful, Von Neumann and Morgenstern (1953) (p.19) have already criticized the validity of complete preferences from a positive standpoint:

> It is conceivable – and may even in a way be more realistic – to allow for cases where the individual is neither able to state which of the two alternatives he prefers nor that they are equally desirable.

Aumann (1962) and Bewley (2002) point out that when choices are very hypothetical or very complex, it is not clear that an individual should have a definite preference over the alternatives. Bewley (2002) and Ok (2002) are at the origin of representation theorems of incomplete preferences, in various contexts: certain, risky and ambiguous domains. A more thorough review is available in Appendix D.

There are many models of incomplete preferences in the literature, but there are comparatively very few studies on their empirical validity. Studying incomplete preferences is challenging, and to the best of my knowledge, only five experiments have investigated completeness, more or less directly: Danan and Ziegelmeyer (2006), Cettolin (2016), Sautua (2017), Qiu and Ong (2017) and Costa-Gomes et al. (2016). This small number should be contrasted with the large body of literature on completeness in psychology and management sciences (see Deparis (2013) for example), as well as the large body of evidence on violations of transitivity in economic contexts.

The main reason for the relative lack of experiment on the completeness axiom is conceptual. Jointly, revealed preferences and singleton choice forbid the exploration of the completeness axiom, as explained by Mandler (2005). First, assuming revealed preferences, an observer infers that chosen alternatives are preferred to unchosen ones. Second, singleton choice forces decision makers to choose precisely one alternative in the choice set. These two assumptions together impose what Mandler has called "observational completeness". By definition, this setting assumes right away that chosen alternatives are comparable with unchosen ones, that is, preferences are complete. I solve this problem by using choice correspondences instead of choices functions, in contrast with all other experiments which use indirect mechanisms, except Costa-Gomes et al. (2016).

Eliaz and Ok (2006) introduces a weakening of WARP called the Weak Axiom of Revealed Non Inferiority (WARNI hereafter), with the interpretation that chosen alternatives are *not worse* than any other available alternative. This interpretation does *not* assume completeness of preferences anymore.

Axiom 4.2 (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.

$$\text{for all } S \in 2^X, y \in S, \text{ if for all } x \in c(S), \text{there exists a } T \in 2^X, y \in c(T), x \in T \Rightarrow y \in c(S)$$

WARNI states that if an alternative is not worse than all chosen alternatives, it must be chosen too. A choice correspondence satisfies WARNI if and only if it is rationalized by a reflexive and transitive strict preference relation – i.e., a partial order. Aleskerov, Bouyssou, and Monjardet (2007) give an equivalent axiomatization of rationalizability by a partial order, which is given in Appendix D.3.[21] Their axiomatization is linked to Sen (1971)'s decomposition of WARP, rather than a direct weakening of WARP.

4.2.2 Conditions for Rationalizability by Intransitive Indifference

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. The famous "coffee and sugar" example of Luce (1956) illustrates it.

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. Here, I will test two models, the original semi-order model of Luce (1956) and the interval order model of Fishburn (1970). Aleskerov, Bouyssou, and Monjardet (2007)'s Chapter 3 provide testable conditions for rationalizability by a semi-order and an interval order.

4.2.2.1 Interval order

An interval order is a partial order that satisfies the strong intervality axiom.

Definition 4.1 (Strong Intervality (Ferrers property))**.** A binary relation $\succ$ satisfies Strong Intervality if and only if x is strictly better than y and z is strictly better than t, then either x is strictly better than t or z is strictly better than y. Formally:

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

[21] I do not formally prove this equivalence. Theorem 2 of Eliaz and 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.

Interval orders are representable by a utility function u and a threshold function t depending on the worst alternative. That is:

$$\text{for all } x, y \in X, x \succ y \Leftrightarrow u(x) - u(y) > t(y)$$

Interval orders are not an entirely accurate representation of the Weber-Fechner law, as the threshold does not depend on the difference between the two alternatives, but only on one of the alternative. 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 4.3 (Functional Asymmetry (FA))**.** A choice correspondence satisfies *Functional Asymmetry* if

$$\text{for all } S, S' \in 2^X, c(S) \cap (S' \backslash c(S')) \neq \emptyset \Rightarrow c(S') \cap (S \backslash c(S)) = \emptyset$$

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.

4.2.2.2 Semi-order

A semi-order is an interval order that satisfies the semi-transitivity axiom.

Definition 4.2 (Semi-transitivity)**.** A binary relation $\succ$ satisfies semi-transitivity if and only if x is strictly better than y and y is strictly better than z, either x is strictly better than t or t is strictly greater than z. Formally:

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

Every semi-order is representable by a utility function u and a constant threshold t. That is:

$$\text{for all } x, y \in X, x \succ y \Leftrightarrow u(x) - u(y) > t$$

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 4.4 (Jamison-Lau-Fishburn (JLF)). A choice correspondence satisfies the *Jamison-Lau-Fishburn* axiom if

$$\text{for all } S, S', S'' \in 2^X, S \subseteq (S' \backslash c(S')), c(S') \cap S'' \neq \emptyset \Rightarrow c(S'') \cap (S \backslash c(S)) = \emptyset$$

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).

Interval orders and semi-orders are strengthening of partial orders and weakening of classical preferences. No direct experiment on intransitive indifference models has been made in economics, despite the large body of literature in psycho-physics.[22]

4.3 Incomplete and Intransitive Indifference Preferences in Practice

Table 7 shows how the three models above rationalize ε-correspondences. Using weakening of classical preferences is not very helpful to rationalize ε-correspondences in my experiment. The various weakenings rationalize marginally more ε-correspondences than classical preferences, and the difference is never significant. Interestingly, partial orders, interval orders, and semi-orders rationalize the same ε-correspondences in my experiment. It does not have to be the case in theory. Failures of completeness here can be understood as a failure of transitivity of the indifference, in line with Eliaz and Ok (2006)'s reinterpretation of intransitive indifference as incompleteness.[23] I decompose the results by additional payments and information in Appendix C.2. The picture is the same: incomplete preferences and intransitive indifference do not rationalize many more choice correspondences.

Table 7: Rationalizability by a weakening of classical preferences.

	ε-correspondence	Choice correspondence
Classical preferences (WARP)	46%	98%
Semi-order	49%	98%
Interval order	49%	98%
Partial order (WARNI)	48%	98%

In my experiment, contrary to most results elsewhere in the literature, relaxing completeness does

[22]Sautua (2017) rules out intransitive indifference as explaining his observations, but it is not a test of intransitive indifference models *per se*.

[23]This comes from how indifference and incompleteness are distinguished in their rationalization of observations. Preferences are assessed from binary sets. If $c(\{x,y\}) = \{x,y\}$, then x and y are either indifferent or incomparable. If they are incomparable, then there must exist a third alternative z which is incomparable with one of them and strictly ranked with the other. This is close to the semi-transitivity axiom.

not rationalize more choice correspondences. In Danan and Ziegelmeyer (2006), 59% of subjects exhibit some incompleteness in the choice of risky lotteries. In Cettolin (2016), around half of the subjects exhibit some incompleteness in the choice of ambiguous lotteries. In Sautua (2017), many subjects exhibit some incompleteness in the choice of ambiguous lottery tickets. In Qiu and Ong (2017), 92% of subjects exhibit incompleteness in a strategic environment. In Costa-Gomes et al. (2016), 5% of their subjects can be rationalized by incomplete preferences, and 21% are the closest to having incomplete preferences. This last experiment is the closest to mine. The most significant difference is the incentive to choose several alternatives. In their experiment, subjects can postpone, with a cost, and choose again later. In that case, subjects can get more information on the alternatives (headphones).

The relatively low proportion of subjects with incomplete preferences in the last experiment is close to my result. One explanation for the difference with the other four experiments is the kind of alternatives we have used. Costa-Gomes et al. (2016) and myself explore the certain domain, not the risky/ambiguous one. A more detailed explanation of the difference between the experiments is available in Appendix D.2.

4.4 Menu-Dependent Choice

The results of the previous subsections imply that weakening transitivity or completeness while keeping menu-independence does not rationalize ε-correspondences. Sen (1997) explained why in some settings it is sensible to go beyond the traditional menu-independent maximization paradigm. The elicitation of ε-correspondences allows me to study quite a wide range of menu-dependent maximization models (Aleskerov, Bouyssou, and Monjardet (2007), Frick (2016), and Tyson (2018), among others) for the first time. In the following, I concentrate on three models, one by Frick (2016) and the other two by Aleskerov, Bouyssou, and Monjardet (2007).

4.4.1 General Threshold Models

The menu-dependent maximization models I study with choice correspondences extend just-noticeable differences. The main difference is the introduction of a threshold *function*, which depends on the menu and the alternatives. Generally, a threshold model has the following representation.

Definition 4.3 (General Threshold Representation)**.** A choice correspondence c on X admits a threshold representation if there exist functions $u : X \to R$ and $t : X \times X \times 2^X \to \mathbb{R}^+$ such that for every S,

$$c(S) = \left\{ x \in S \,\middle|\, \max_{y \in S} u(y) - u(x) \leq t(x, y, S) \right\}$$

u is the fully rational benchmark and t the departure threshold of the representation. 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 can capture attraction or decoy effects for instance.

Threshold representations studied here depend on the choice set and sometimes on the alternatives considered. Thresholds can depend on various combinations of x, y, and S, such as $t(x, S)$ for instance. Aleskerov, Bouyssou, and Monjardet (2007) showed that these threshold representations reduce to menu-dependent ($t(S)$) or context-dependent ($t(x, y, S)$) models.[24] The unique utility allows me to recover a unique preference for each decision maker. It is one advantage of using a variation of intransitive indifference rather than a variation of incompleteness. Incomplete preferences are represented by multi-utility models, which yield several conflicting orders of alternatives.

4.4.1.1 Monotone Threshold (MT)

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 $S' \subseteq S$. This is one way to take into account choice overload. When the size of the choice set is larger, the choice is less precise, and the threshold is larger. Frick (2016) shows that *occasional optimality* characterizes the *monotone threshold* on choice correspondences.

Axiom 4.5 (Occasional Optimality)**.** A choice correspondence satisfies occasional optimality if, for all $S \in 2^X$, 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 makers chooses from S is optimal. Occasional optimality requires that at least some of the decision maker's choices from S be optimal.

4.4.1.2 Menu-Dependent (MD) 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)

[24] More 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. I have decided to keep 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. I have decided to keep the latter representation.

(p.154-155) provide the testable condition for a menu-dependent threshold model to rationalize a choice correspondence.

Definition 4.4 (Strict Cycle of Observation)**.** A *strict cycle of observation* are n sets $S_1, S_2, \ldots, S_n$ in 2^X such that:

$$(S_1 \backslash c(S_1)) \cap c(S_2) \neq \emptyset$$
$$(S_2 \backslash c(S_2)) \cap c(S_3) \neq \emptyset$$
$$...$$
$$(S_n \backslash c(S_n)) \cap c(S_1) \neq \emptyset$$

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

Axiom 4.6 (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 who satisfy functional acyclicity, I only directly elicit strict preferences, not indifference. Strict Preferences are acyclic and therefore can be augmented using the transitive closure. It is not possible to complete the preferences of the decision maker, however. It is in particular not possible to assume that the remaining binary relations are indifference relations, as shown in Appendix C.4.

4.4.1.3 Context-Dependent (CD) 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. Aleskerov, Bouyssou, and Monjardet (2007) (p.156-157) give the condition on choice correspondences for context-dependent rationalizability.

Axiom 4.7 (Fixed Point)**.** A choice correspondence c satisfies *fixed point* if for any $S \in 2^X$, 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 Sen's α. α requires that all alternatives chosen in a set be chosen in subsets. Fixed point only requires that one alternative chosen in a set be 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 is 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. Now take the set of fixed points in $X \backslash FP(X)$, it is the most preferred alternatives in $X \backslash FP(X)$, 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 preferences are complete. Compared to strict revealed preferences, this revelation of preferences also reveals indifference.

4.5 Menu Effects in Practice

Table 8 shows that generalized threshold representations rationalize significantly more ε-correspondences than classical preferences when they violate WARP. This result holds when the choice correspondence is partially or fully identified. Looking at the all ε-correspondences in Table 9 nuances the results. When I partially or fully identify the choice correspondences, the decision makers satisfy WARP more often, and the significance of the differences decreases. From both tables, it is clear that introducing menu effects increases the proportion of ε-correspondences for which welfare assessments are possible.

When comparing the different models, the menu-dependent model explains significantly more than the monotone threshold one. The monotonicity of the threshold is a real constraint on rationalizability. The context-dependent model rationalizes a bit more than the menu-dependent one. The difference between the two is significant when I restrict to the sample where ε-correspondences violate WARP. When looking at the explanatory power of the two models in Appendix C.3.1, the context-dependent model has five times less power than the menu-dependent one. The large rationalizability by menu-dependence in Table 9 means that most strict preferences are consistent, even when ε-correspondences violate WARP.

Table 8: Comparison of different models of threshold rationalization to semi-order-rationalizability when WARP is violated. Significance levels are with semi-order rationalizability. Note that menu-dependent and context-dependent are only significantly different from each other at the 1% level on ε-correspondences.

	ε-correspondence	Partial	Choice correspondence
Semi-order	4%	2%	0%
Monotone threshold	26%***	25%***	0%
Menu-dependent	77%***	74%***	100%***

	ε-correspondence	Partial	Choice correspondence
Context-dependent	88%***	79%***	100%***

Table 9: Comparison of different threshold rationalizability to semi-order and weak order rationalizability. Significance levels are with semi-order rationalizability. Note that menu-dependent and context-dependent are significantly different from each other at the 1% level on ε-correspondences, but not when they are partially identified.

	ε-correspondence	Partial	Choice correspondence
Classical preferences	46%	59%	98%
Semi-order	49%	59%	98%
Monotone threshold	60%***	69%	98%
Menu-dependent	88%***	89%***	100%
Context-dependent	93%***	91%***	100%

I compare in Table 10 the different models depending on the information provided. Subjects are on average more consistent in the sentence treatment, because they express more indifference, as I will show in Section 5. This trend goes on when WARP fails. The training treatment, where subjects express the most strict preferences, reaches similar satisfaction rates in the menu-dependent and the context-dependent models. The video treatment lags behind the other two with these last two models, and the differences are significant. The relative increase in rationalization in the sentence and training treatments imply that strict revealed preferences are more consistent in these two treatments. In the sentence treatment, there are relatively few of them, which means they are more easily acyclic. In the training treatment, the decision makers have a more precise idea of the task they want or do not want to perform and choose more consistently as a consequence.

Table 10: The different threshold models by the information provided. The significance levels are with respect to the video treatment.

	Sentence	Video	Training
Classical preference	62%***	44%	45%
Semi-order	65%***	46%	46%
Monotone threshold	75%***	56%	61%
Menu-dependent	94%***	81%	93%***
Context-dependent	98%***	86%	97%***

5 Indifference

One reason to identify choice correspondences is to study the prevalence of indifference relations. Indifference between alternatives might generate failures of WARP even when the *as if* preference is reflexive, complete and transitive. Indifference here will be assessed using two different methods. First, for ε-correspondences which satisfy WARP, it is revealed indifference as given at the beginning of Section 4. For ε-correspondences who do not satisfy WARP, but satisfy fixed point, it is revealed as explained in Section 4.4.1.3. I cannot assess indifference from choice functions using the same assumption on the revelation of preferences.

I relax this assumption for choice functions, assuming now that chosen alternatives are *weakly* better than unchosen ones. With ε-correspondences, it would contradict the idea that pay-for-certainty elicit all maximal alternatives. On choice functions, however, it might be a more sensible assumption, and it allows the study of indifference. In that case, a choice reversal is revealing indifference, rather than inconsistency. Using weak revealed preferences for choice functions increase the satisfaction of WARP to 80%, while the power is still acceptable.

I show in Table 11 how the preference is divided between strict relations and indifference relations for each subject, on average. I show only results when WARP is satisfied for choice correspondences, as it keeps almost the whole subsample of choice correspondences. There are six possible comparisons between alternatives. A classical decision, therefore, has six binary relations in total. Depending on the sample considered, the precise division between indifference and strict preferences vary. One thing is clear, however, choice correspondences and ε-correspondences show a lot more indifference than choice functions.[25]

Table 11: Indifference and strict preferences, for choice functions, correspondences, and ε-correspondence.

Axiom	Choice	Indifference	Strict Preference
WARP	Function	0.29***	5.71***
WARP	Correspondence	2.96	3.04
WARP	ε-correspondence	3.00	3.00
FP	ε-correspondence	2.15***	3.85***

In Table 12 and 13, I show how much indifference relations there are on average for each subject, by information provided and additional payments. The more information provided, the less indifference there is. The more additional gain there is to select an alternative, the more indifference there is. When comparing choice functions and choice correspondences with the same treatment, choice

[25] The p-value of the difference in the number of indifference relation between choice function and choice correspondences is 10^{-12}.

functions show significantly less indifference.

Table 12: Average number of indifference relations with ε-correspondences and choice correspondences, by information provided. Significance levels are with choice correspondences which satisfy WARP.

Axiom	Choice	Sentence	Video	Training
WARP	Function		0.29***	
WARP	Correspondence	4.33	2.61	2.73
WARP	ε-correspondence	3.98	2.81	2.36
FP	ε-correspondence	3.53***	1.36***	1.07**

Table 13: Average number of indifference relations with ε-correspondences, by additional payment. The average number of indifference relations within the same payment, between subjects who satisfy WARP or FP are significantly different from each other. The average number of indifference relations within the subjects who satisfy WARP or FP, between additional payments are significantly different from each other when one of them is a no additional payment.

Axiom	No	Low	High
WARP	2.09	3.43	3.54
FP	0.91	1.52	1.45

Indifference is not homogeneous among all subjects. In Figure 2, I show that some subjects are fully indifferent, whereas some have only strict preference relations. The evolution between the level of information provided is the same as the average one, however. The more information is provided, the less indifferent subjects are.

Indifference is widespread with choice correspondences. It is significantly higher than what could have been obtained with choice functions, even assuming that revealed preferences are weak. It is also robust, as it does not disappear even in the training treatment, where subjects had quite complete information about the tasks. The additional payments also have an impact on the indifference, but even with no payments, indifference relations appear. Even a simple elicitation method for choice correspondence, which does not identify the choice correspondence alone, is better than choice functions to study indifference.

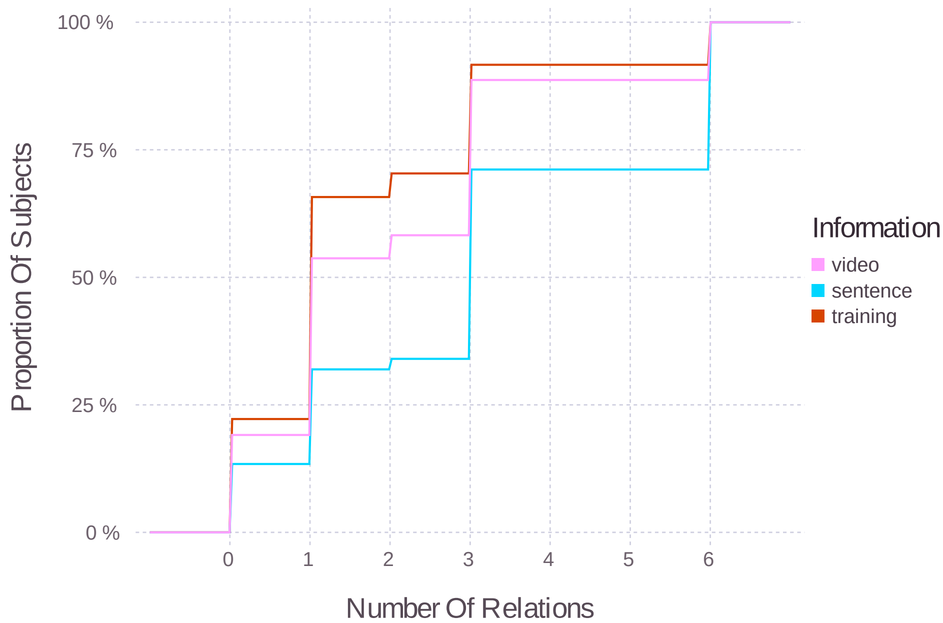

Figure 2: CDF of the number of indifference relations, by information provided. The p-value of the Mann-Whitney test of the difference between the training and the video treatment is 0.055. For the sentence and video treatment, it is 0.0001

6 Aggregating Choice Correspondences

Finally, identifying choice correspondences might be a way to aggregate decision makers' choices or preferences more easily. My experiment is not about a social choice nor explicitly aimed at collective choices. Choice correspondences allow me to recover larger sets of Pareto-superior alternatives for each decision maker. These larger sets of Pareto-superior alternatives should intersect more often, and thus should agree more often on which alternatives are the best. This section is the first comparison between choice correspondences and choice functions in that regard. I take two approaches to do so.

6.1 Condorcet

The first approach that is close to voting procedure, namely, looking at the Condorcet winner for each choice correspondence.

Definition 6.1 (Condorcet winner)**.** An alternative x is a *Condorcet winner* for a correspondence c if it is chosen in all binary choices with other alternatives from X.

$$\text{for all } y \in X, y \neq x, x \in c(\{x, y\})$$

Note that with choice functions, there can be only one Condorcet winner per subject. With choice correspondences, however, all alternatives might be Condorcet winners. This approach is 'model-free', in the sense that it does not require any rationale for the choice. It is a rather permissive approach in theory, as for having at least one Condorcet winner, subjects only have to be consistent within binary choices, not within the full powerset. That is, I do not take into account choices in sets of size three or four.

Once I have the Condorcet winners for each choice correspondence or function, I can aggregate them. I use here a procedure that is close to approval voting.[26] That is, I assume that decision makers would vote for all their Condorcet winners.

6.2 Maximal Alternatives

I have built the second approach on the revealed preferences. When subjects satisfy WARP, I use revealed preferences, and when they satisfy fixed point, the corresponding preferences. I drop the rest of the sample of ε-correspondences (7% of the pool) and choice functions (43% of the pool).[27]

I again consider approval voting, where subjects would vote for all their maximal alternatives. Note that here, all preferences must have at least one maximal alternatives. For choice functions, they only have one, but for choice correspondences, potentially all alternatives are maximal.

6.3 Aggregation in Practice

Tables 14 and 15 are similar in the lessons I can draw from them. The number of maximal alternatives or winners is much higher with choice correspondences or ε-correspondences than it is with choice functions. Importantly, with correspondences, at least one alternative is approved by the majority of the voters. The lower approval for ε-correspondences compared to choice correspondences is a consequence of the lower number of indifference relations when ε-correspondences satisfy fixed point, as shown in Section 5.

[26] See Laslier and Sanver (2010) for a more thorough overview of the approval voting literature.

[27] I could infer preferences for these subjects, using, for instance, behavioral welfare economics, as done in Bouacida and Martin (2017), but in a sense assessing Condorcet winners already does so.

Table 14: Which alternative is the Condorcet winner? What is the corresponding aggregate preference? Average of each alternative being a winner, depending on the choice procedure considered. Significance levels are with respect to choice correspondences.

	Addition	Spellcheck	Memory	Copy	Preference
Choice function	29%***	24%***	22%***	24%***	$A \succ S \sim C \succ M$
Choice correspondence	73%	73%	65%	69%	$A \sim S \succ C \succ M$
ε-correspondence	65%	56%	58%**	68%	$C \succ A \succ M \succ S$

Table 15: Which alternative is maximal? Average of each alternative being maximal, depending on the choice procedure considered. Significance levels are given with respect to choice correspondences. ε-correspondences, when they satisfy WARP or FP, are significantly different from each other.

Axiom	Choice	Addition	Spellcheck	Memory	Copy
WARP	Function	24%***	24%***	21%***	31%***
WARP	Correspondence	73%	73%	65%	69%
WARP	ε-correspondence	70%	68%	65%	74%
FP	ε-correspondence	48%***	40%***	44%***	52%**

I can also look at the aggregate order of preferences provided by these two aggregation methods. Choice functions and choice correspondences do not rank the same alternative as the best, using inferred preferences. Addition and spellcheck are the best on average with choice correspondence, whereas copy is the best with choice functions. With Condorcet winners, on the other hand, both agree that addition is one of the best tasks.

In Table 16, I look at the influence of the information provided on the aggregation, using Condorcet winners. I consider only ε-correspondences for clarity, but the picture is similar to choice correspondences. The sample is just becoming very small in the latter case. The trend is similar when looking at maximal alternatives. Some trends appear clearly. The information given decreases the proportion of subjects who would choose the memory task, more than the average decrease due to the increase in quantity and quality of information and the sharpening of preferences it implies. The copy task, on the other hand, remains chosen at around the same levels by everyone.

The value of information is positive on average in my experiment, something that also appears in the average gains in the tasks depending on the treatments. They went from €2.66 in the sentence treatment to €3.23 in the video treatment to reach €3.88 in the training treatment.

Table 16: Proportion of ε-correspondences for which a given alternative is a Condorcet winner, depending on the information provided. Significance levels are given for a task, between the treatments.

Treatment	Addition	Spellcheck	Memory	Copy
Sentence	72%	62%*	85%***	75%*
Video	64%	51%	58%	65%
Training	59%	65%***	36%***	72%

Overall, using choice correspondence allow for more agreement across decision makers.

7 Conclusion

In this paper, I have introduced a new method to identify choice correspondences, *pay-for-certainty*. I fully identify the choice correspondence of 26% of the subjects in my experiment. One natural question is how to raise that identification rate. Decreasing the value of ε seems out of the question in this experiment. Another way is to increase the stake, but it is not always possible.

One way to improve the identification of a choice correspondence is to change the incentive to choose all maximal alternatives. One possibility is to use as an incentive a small probability of getting several alternatives. If the decision maker chooses x and y in the set $\{x, y\}$, she gets both with a small probability (1%), and the remaining probability is allocated uniformly between x and y. One potential problem of this mechanism is the well-known sensitivity of subjects to small probabilities.

The pay-for-certainty method is suited for non-risky, non-ambiguous alternatives. Extending the method to other kinds of alternatives requires additional assumptions. The adaptation to the Machina-Marschak triangle à la Andreoni-Sprenger used for studying behaviors in risky environments is not straightforward. The current mechanism allows decision makers to create their lottery. Additional research should be done in this direction in order to understand how to adapt the method in the infinite case. I give some preliminary directions in Appendix A.2.

Even if limited to certain kinds of alternatives, eliciting a choice correspondence has many benefits. I have been able to study rationalization of observed choices when subjects do not behave *as if* they have transitive, reflexive and complete preferences. I could relax transitivity and completeness, although this did not explain significantly more data. The likely explanation for the low occurrence of incomplete preferences is in the kind of alternatives I have considered. Sure outcomes did not yield incomplete preferences in Costa-Gomes et al. (2016) as well. Introducing menu effects is much more promising. The context-dependent model rationalizes 93% of observed choice correspondences. Importantly, choice correspondences allow the assessment of preferences that choice functions do not. I show that indifference is significant in my experiment. By assessing indifference, choice correspondences enlarge the set of Pareto-superior alternatives. It implies that in practice, reaching an agreement across subjects is easier than generally believed.

Finally, the context-dependent model is the most successful model in my experiment. One limit, however, is that the choice sets were rather small. Considering that choice processes might limit the number of alternatives considered, for various reasons (see Caplin, Dean, and Martin (2011), Manzini and Mariotti (2012) for instance), it is an open question how the context-dependent model rationalize choices in larger sets. It is a challenging question to tackle. Above sets of five alternatives, it is no longer realistically possible to study the whole powerset, and thus to falsify the model. Nevertheless, the results in the experiment show that choices depend on the choice sets, even if preferences do not. In that sense, it is consistent with the literature on reference-dependence and status quo bias in behavioral economics. The fixed point in the whole set is the reference. It provides a revealed preference based method for studying these phenomena, in a similar spirit as Ok, Ortoleva, and Riella (2015).

A Proofs and Extension of Pay-For-Certainty

A.1 Proof of Proposition 1.1

Proof. Take $S \in 2^X$, $\varepsilon < \varepsilon'$, I will show that any $x \in S \backslash c_{\varepsilon'}(S)$ is also in $S \backslash c_{\varepsilon}(S)$. Take $x \in S \backslash c_{\varepsilon'}(S)$:

- By constrained maximization of the payments:

$$x \in S \backslash c_{\varepsilon'}(S) \Leftrightarrow \text{ there exists } y \in S, (y, 0) \succ_2 (x, \varepsilon')$$

- By the remark in Section 1.4, I have that y also ε-dominates x.
- By constrained maximization of the payments again, $x \in S \backslash c_{\varepsilon}(S)$.
- $$S \backslash c_{\varepsilon'}(S) \subseteq S \backslash c_{\varepsilon}(S) \Leftrightarrow c_{\varepsilon}(S) \subseteq c_{\varepsilon'}(S)$$

  .

□

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

I discuss here why extending to X 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.

A.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.

A.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. I can expand on the previous example, with the following modification: the choice set is {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 her certainty equivalent is 0.41? Assuming that ε is small enough, would she choose 0.4 or {0.4, 0.5}? It is not clear. If she chooses 0.4, what is her certainty equivalent? I certainly cannot guarantee that it is equal to 0.4.

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.

A.3 Selection Mechanisms

A.3.1 Influence of the Uniform Selection Mechanism

The selection mechanism might influence which alternatives decision makers choose. To understand why I first explain the reason behind choosing a non-singleton set of alternatives. One reason to choose two alternatives or more is when one is indifferent between them. If a decision maker is indifferent between x and y, she might not care if I tell her that she will get one of them were she to choose both. Another reason for choosing more than one alternative is indecision. A decision maker might not know which alternative she prefers. Indecision usually arises from a lack of information about the alternatives and how to compare them. Imagine a decision maker arriving in a new country who gets offered two different fruits she did not know before. Which one to choose? She probably wants to get more information about them before choosing, both to learn about the alternatives and to learn about her tastes. In experiments, subjects are rarely allowed to experience the alternatives before choosing them. Costa-Gomes et al. (2016) is a rare instance allowing this, and decision makers postponed their choice in order to get more knowledge.

In the context of a choice correspondence elicitation, still, another reason to choose more than one alternative might be the selection mechanism. Decision maker might take advantage of it to build new alternatives. Think about the uniform selection mechanism and the following alternatives in X:

- Lottery a: 50 cents for sure
- Lottery b: \$1 with probability 1/2 and \$0 with probability 1/2.

With the pay-for-certainty procedure and a classical expected utility maximizer decision maker, observed choices should be:

- Risk-averse decision makers choose $\{a\}$;
- Risk-loving decision makers choose $\{b\}$;
- Risk-neutral decision makers choose $\{a, b\}$.

Pay-for-certainty identifies as risk-neutral slightly risk-loving and slightly risk-averse decision maker if ε is small enough. With the uniform selection mechanism, however, choosing $\{a, b\}$ creates a new lottery:

- Lottery c: 50 cents with probability 1/2, \$1 with probability 1/4 and \$0 with probability 1/4.

Lottery b is a mean-preserving spread of lottery c, and lottery b is a mean-preserving spread of lottery a. For expected utility decision makers, lottery c does not matter, as it is in between a and b regarding the risk involved. When decision makers are not expected utility maximizers, it might be a problem. Now, the uniform selection mechanism allows the decision maker to create a "new" alternative. She might prefer this "new" alternative to the two original ones. Importantly, it is not possible looking at choices whether the choice of $\{a, b\}$ is driven by risk neutrality or by the will to create the "new" alternative.

One assumption from the literature (see Dekel (1986) and Chew (1989)) guarantees it is not a problem, *betweenness*.

Assumption A.1 (Betweenness)**.** *If the decision maker is indifferent between two alternatives, she is indifferent between any convex combination of both alternatives and the original alternatives.*

$$\text{for all } x, y \in X, x \sim y \Leftrightarrow \text{for all } p \in [0,1], x \sim (1-p)x + py \sim y$$

Betweenness is often violated in settings involving lotteries (see Camerer and Ho (1994) for instance), in particular by decision makers with a preference for randomization as in Agranov and Ortoleva (2017). Pay-for-certainty with the uniform selection mechanism is thus generally not suited for choices between lotteries, but other selection mechanisms can obviate this problem.

One can think of allowing subjects to choose the distribution on the set of alternatives, instead of imposing the distribution. The problem generated by the possibility of creating a new alternative is getting worse here. Another selection mechanism is delegation: someone else makes the selection. The decision maker cannot use the selection mechanism to create new alternatives. It moves the selection process from the risky domain to the ambiguous one, as now the distribution used to make the selection is unknown to the decision maker. The last possibility, *flexibility*, is to allow the decision maker to choose again in the chosen set when it is not a singleton. To be incentive compatible, the additional payments of pay-for-certainty must be costs and not gains. It has been done in experiments by Danan and Ziegelmeyer (2006) and by Costa-Gomes et al. (2016). A more thorough discussion of these mechanisms follows in Appendix A.3.2 and Appendix D.2.

A.3.2 Other Selection Mechanisms

In Section 1.2, I have alluded to other possible selection mechanisms. I discuss here a few of them.

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. Own randomization also requires the betweenness property in order to get the right interpretation of observed choices. It is more complicated to implement experimentally, as it adds more features to the elicitation procedure, but yields to potentially more fruitful results. I have run a pilot with own randomization, on 37 subjects. I have computed their choice correspondence assuming that any alternative with positive probability is a chosen alternative. The sizes of the chosen sets are in Table 17. Only 8% of the subjects choose bayesian distributions, using the distribution over the four alternatives as the 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 the pilot or the own randomization selection mechanism. The details are available by asking the author.

Table 17: Size of the chosen set in the pilot where subjects could choose their distribution over the alternatives.

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

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 the correctness of the 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. Her beliefs on the selection mechanism influence the choices made by the decision. The uniform selection mechanism explicitly states the probabilities of getting an alternative. As long as the decision makers trust the experimenter, reality constrains beliefs. 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. One way to mitigate that might be to delegate anonymously to other subjects in the experiment and not to the experimenter. Beliefs are hard to control here.

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 the literature before, in at least two experiments, one by Danan and Ziegelmeyer (2006) and the other by Costa-Gomes et al. (2016). 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 always optimal. With flexibility, indifference is not properly elicited. 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.

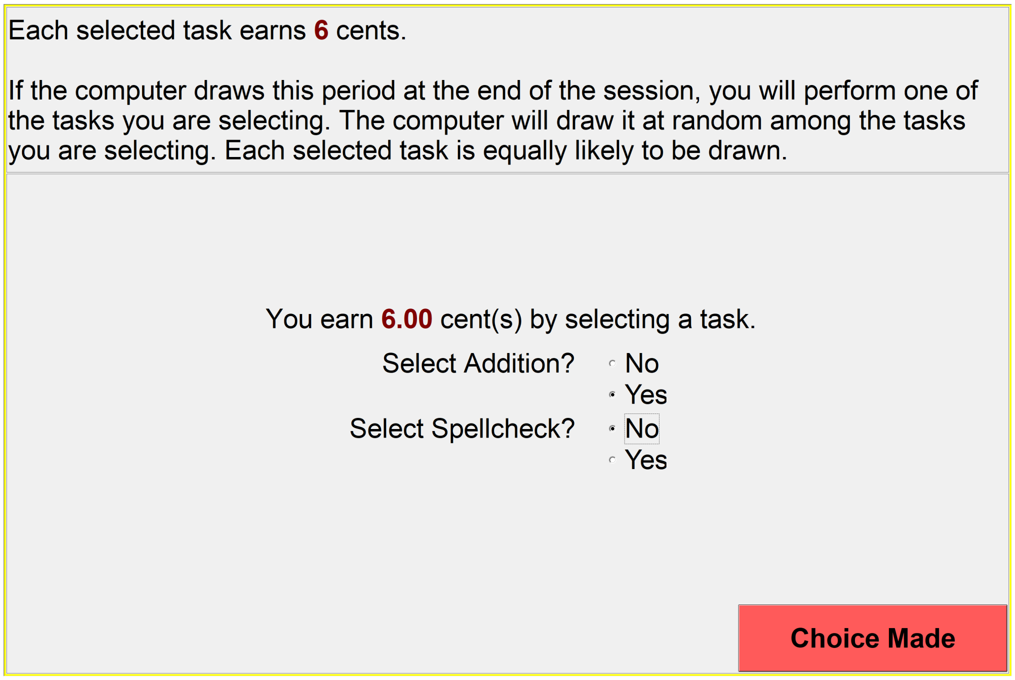

Figure 3: A choice screen.

B Experiment

B.1 Instructions

The experiment happened in France. I gave the instructions in French. The original instructions are available http://www.bouacida.fr/files/eliciting-choice-correspondences/instructions/.

B.2 Screenshots from the Experiment

Each choice with pay-for-certainty happened in two steps. First, subjects faced the choice screen, as in Figure 3, where they said "yes" or "no" to each available alternative. Second, they faced the confirmation screen, as in Figure 4, where they saw a summary of their choice and the associated additional payment. Choices were taken into account only after subjects confirmed them on the confirmation screen.

B.3 Demographics of the Sample

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 5 shows the age distribution of the sample.

Figures 6 and 7 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

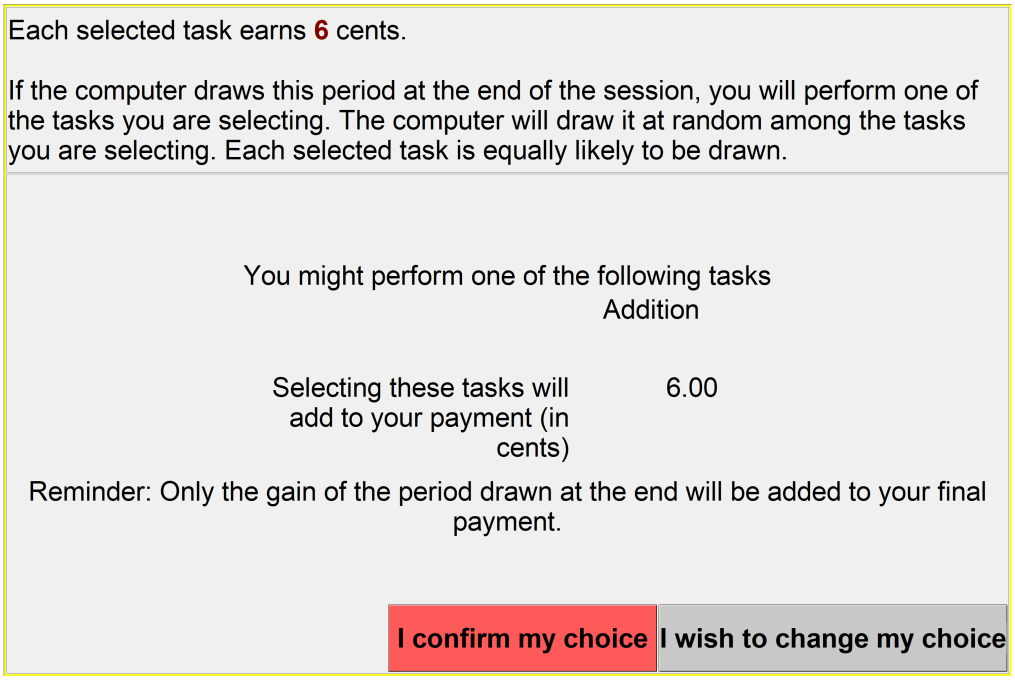

Figure 4: A confirmation screen.

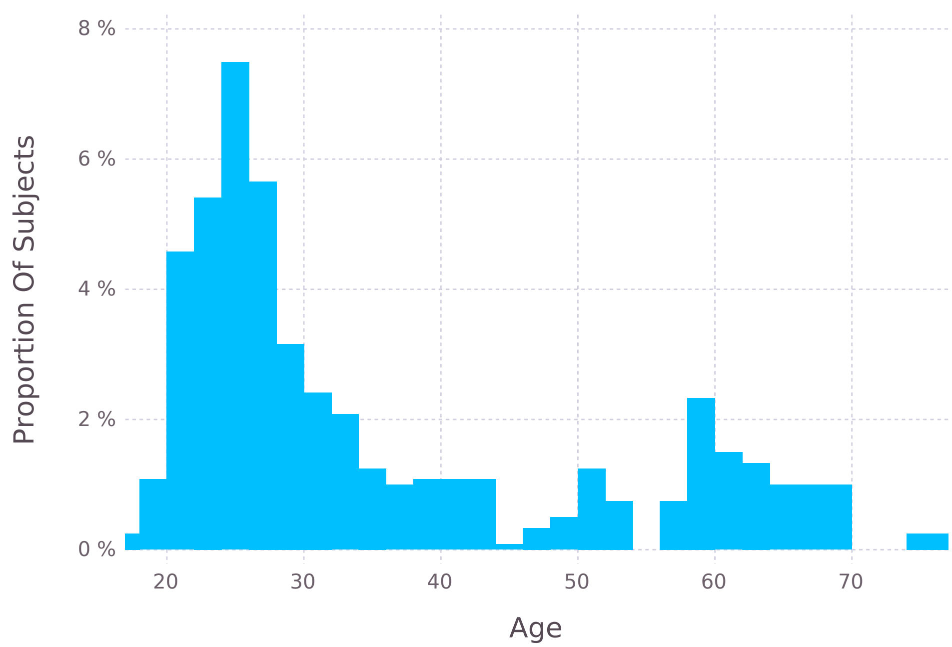

Figure 5: Age distribution of the sample

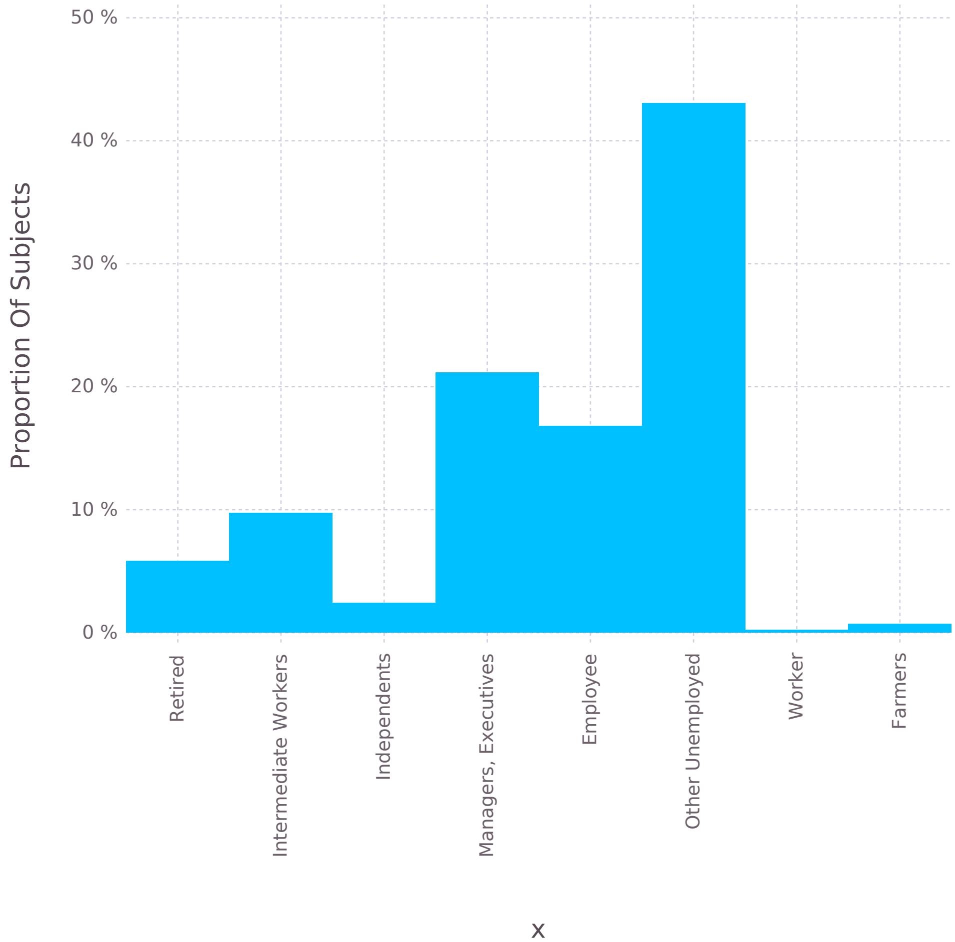

Figure 6: Professional occupation, when given (151 subjects)

French Statistical Institute.

The sample is almost gendered balanced. There are 54.74% of female in the sample. The level of education question only led to a response rate of 42%. Among those who answered, 53.75% have at least started a college education. Overall, the largest population of the sample is a student population.

B.4 Risk Aversion

Figure 8 shows the distribution of the certainty equivalents of the sample. Some subjects must have misunderstood the certainty equivalent elicitation, as a certainty equivalent of 0€ is never optimal. Certainty equivalents above 2.5€ indicate risk loving. Certainty equivalent of 3€ bundles certainty equivalents of 3€ and above.

One metric risk aversion could play a role in is the size of the chosen set. In the experiment, the computer drew at random the task they got when subjects chose more than one task. I could

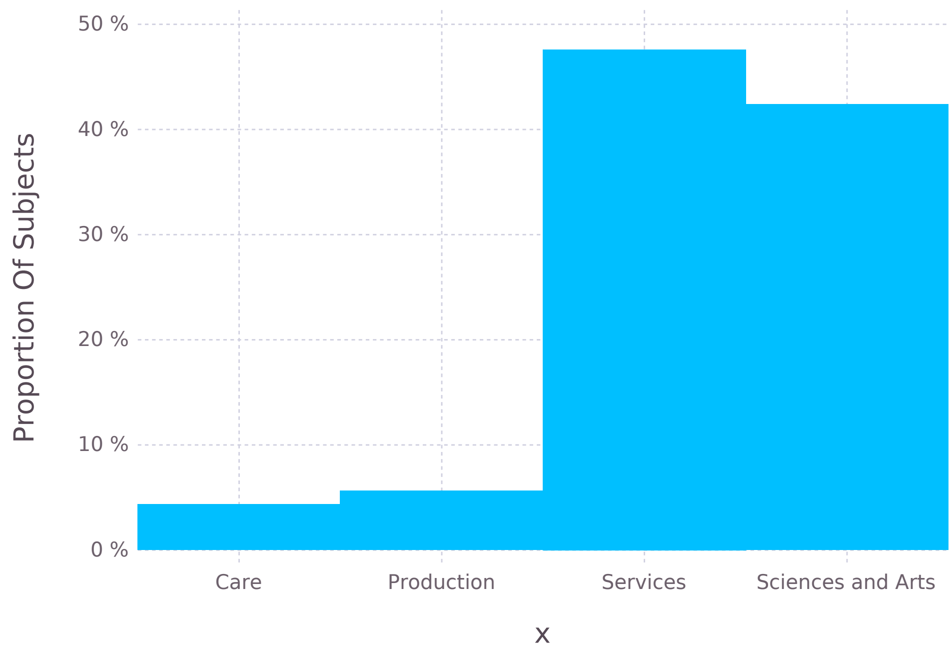

Figure 7: The domain of occupation, when given (172 subjects)

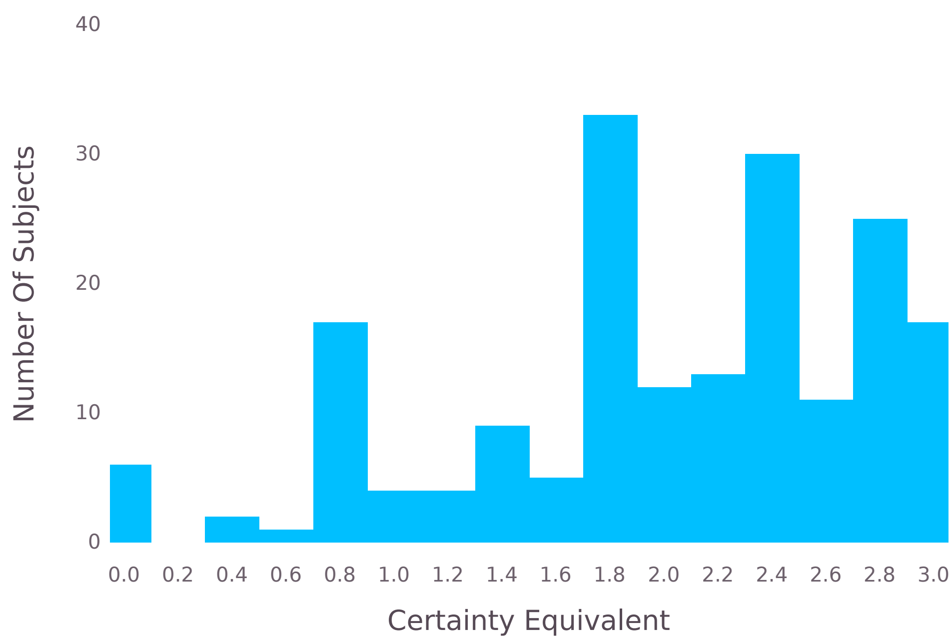

Figure 8: Histogram of the certainty equivalents obtained in the experiment.

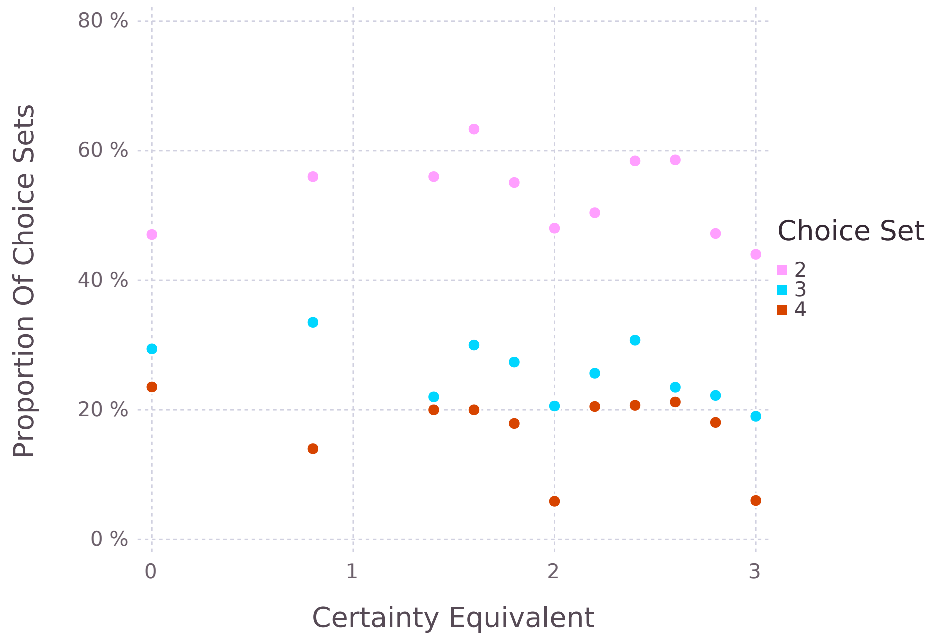

Figure 9: Proportion of singletons depending on the certainty equivalent and the size of the choice set. I kept only certainty equivalents with more than five subjects.

expect people with higher risk aversion to dislike the random draw and choose more singletons. Figure 9 show that it is not the case. If anything, the higher the certainty equivalent, the lower the proportion of singleton is. Some caveat must be given, however: certainty equivalents of 0 are a misunderstanding of the measure of certainty equivalents. I have no explanation for the drop in certainty equivalents at 2€. At a certainty equivalent of 3€, however, it might be because we bundle here 3€ and above. I have computed the correlation numbers between certainty equivalent and the proportion of singletons, but they are highly sensitive to dropping/adding data points and are therefore not very reliable.

C Experimental Results

C.1 Correlation with the Training and the Stated Preferences

In the paper, I have assessed the rationalizability by different preference models of decision theory. I look now at predictors of the choice of the subjects. I do not have much information about the subjects of the experiment, but I have two important ones for evaluating consistency:

1. For the sample in the training treatment, I have their performance during the training for all the tasks.

2. For everyone, I have their answer in the questionnaire. I have asked if there are some tasks they liked or disliked.

These are two proxies for the choices of the subjects expect in the experiment. The training provides them with hard information about their expected payoff. I might not observe it correctly, as subjects might need some time to handle the interface. The answers in the questionnaire are soft information revealed by subjects about their preferences. In both cases, the choice used to select the task is one of the 33 possible. They might get only their third and fourth-ranked alternative in that set.

I have ranked the tasks for each decision maker by their performance. Table 18 shows the ranking of the task performed in the end using the performance in the training sessions. There is a clear correlation between the two. I expected all tasks performed in the end to be third rank or above. It is not the case. Two explanations: the ranking I built is only a noisy representation of their ranking. The cost has driven some subjects to choose all alternatives all the time.

Table 18: Proportion of chosen tasks that is better than the first, second or third best performing alternative during the training, for the sample who trained on the tasks. In case of a tie during the training, tied alternatives are ranked at the highest possible.

	first	second	third
Proportion	32.43%	64.86%	86.49%

I make a second assessment of the training sessions. I build the order corresponding to the observed performance, and compare the choice it induces to the observed choices. Call c_T the choice correspondence induced by the training. If I assume that c_T accurately reflects the choice correspondence of the subjects, I should have that for all S, $c_0(S) \subseteq c_T(S) \subseteq c_\varepsilon(S)$, for any positive ε. Because the order I have built is a weak order, subjects who do not satisfy WARP are unlikely to satisfy the two inclusions. I will, therefore, restrict it to subjects who satisfy WARP. 11% of the subjects satisfy the first inclusion and 52% the second one. These two results imply that the training influences the choices made by the subjects.

I perform a similar analysis with their stated preferences. For all subjects, I build their stated choice correspondence and assume it is their real one. In that case, the stated choice correspondence contains the no gain choice correspondence for 77% of subjects. The stated correspondence is included in the positive gains one for 66% of subjects. Stated correspondences are a predictor of the choices made by the subjects who satisfy WARP.

C.2 Decomposition of Incomplete Preferences

I provide here more results on the relaxation of transitivity and completeness. I decompose the results by additional payments and information provided. I only show the results for completeness, because semi-order and interval orders are strengthening of partial orders. Incompleteness provides an upper bound on intransitive indifference.

C.2.1 By Information

If uncertainty about the value of alternatives influenced incompleteness, one would expect that giving more information about the alternatives to the subjects in the experiments reduces their incompleteness. Table 19 shows a small but insignificant decrease in the prevalence of incomplete preferences when the quantity and quality of information increase. My experiment does not allow me to say anything about the link between incomplete preferences and information, as very little incompleteness shows up in the first place.

Table 19: Rationalizability by complete and incomplete preferences, depending on the information provided.

Treatment	Sentence	Video	Training
complete	62%	44%	45%
incomplete	65%	46%	46%
incomplete but not complete	3%	2%	1%

C.2.2 By Additional Payments

Table 20 shows a small increase in the prevalence of incomplete preferences when the additional payment increases. The difference between the no and high additional payments is significant at the 5% level. The other differences are not significant. Overall, additional payments do not seem to play a role in the prevalence of incomplete preferences.

Table 20: Rationalizability by complete and incomplete preferences, depending on the additional payment.

Gains	No	Low	High
complete	47%	43%	49%
incomplete	48%	44%	54%
incomplete but not complete	1%	2%	5%

C.3 Explanatory Power

C.3.1 Axioms and Rationalizations

With a set of four alternatives, there are 26,254,935 possible choice correspondence and 20,736 possible choice functions. I have tested all the axioms and different kind of rationalizability on each of them. Table 21 gathers the results.[28] 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 axioms. δ bites a little on choice correspondences and does not bite on choice functions, a characteristic shared with β. Lastly, on choice functions, many axioms have the same pass rate and are satisfied by the same choice functions. All models studied here rationalize the same choice functions. Choice correspondences are useful because they allow me to separate the different models.

Table 21: Probability that a decision maker choosing with a uniform random satisfy an axiom / rationalizability model.

Axiom	Correspondence	Function
WARP \ Weak Order	2.86 10^{-4}%	0.12%
Semi-order	6.97 10^{-4}%	0.12%
Interval order	7.88 10^{-4}%	0.12%
WARNI \ Partial order	8.34 10^{-4}%	0.12%
Occasional optimality \ MT	7.28 10^{-2}%	0.12%
Functional acyclicity \ MD	1.44%	0.12%
Fixed point \ CD	7.27%	0.12%
α	5.68 10^{-2}%	0.12%
β	2.63%	100%
δ	27.69%	100%
γ	1.08%	1.62%
Pairwise transitivity	10.29%	37.5%
Bottom-up rationality	2.60%	1.39%
Top-down rationality	0.13%	0.12%
β'	5.08%	4.86%
Outcast	0.27%	0.12%
Functional asymmetry	1.51%	0.12%
Jamison-Lau-Fishburn	14.57%	1.04%
WARP with weak RP	13.89%	72.93%

[28]The procedure I 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 chosen set at random, using a uniform distribution. It is also *as if* the decision maker decided in each choice set and for each alternative with a coin toss to choose it.

C.4 Completing the Revealed Strict Preference

It is not possible to uniquely complete the strict revealed preferences of a decision maker who satisfies Functional Acyclicity. Imagine a decision maker with the following choice correspondence on $X = 1, 2, 3, 4$ (chosen alternatives are in bold): {1, **2**}, {1, **3**}, {1, **4**}, {**2, 3**}, {2, **4**}, {**3, 4**}, {1, **2, 3**}, {1, **2, 4**}, {1, **3, 4**}, {**2, 3, 4**}, {1, **2, 3, 4**} . The corresponding strict revealed preferences are: $2P1, 3P1, 4P1$ and $4P2$, and revealed indifference is $2I3, 3I4$ and $2I4$. This decision maker does not satisfy WARP, but she satisfies FA, the strict part of revealed preferences is acyclic. Completing the strict revealed preferences with indifference (i.e., $3I2$ and $3I4$) yield cyclic preferences ($4P2, 2I3$ and $3I4$). There are two maximal acyclic completions in this case: adding only $3I2$ or only $3I4$, but there is no reason to pick one over the other. This example shows two important things:

- It is not possible to guarantee, reflexive, transitive and complete preferences when functional acyclicity is satisfied.
- There is no unique transitive completion when functional acyclicity is satisfied.

D Incomplete Preferences

D.1 Theory on Incomplete Preferences

Von Neumann and Morgenstern (1953) (p.29) had already an idea on how to represent incomplete preferences:

> It leads to what may be described as a many-dimensional vector concept of utility. This is a more complicated and less satisfactory set-up, but we do not propose to treat it systematically here.

The systematic exploration of this intuition was first carried out by Bewley (2002) in an Anscombe and R. J. Aumann (1963)'s framework and by series of paper by Ok and coauthors, following Ok (2002). Investigating all models representing incomplete preferences is tedious and is not the object of this paper. The various representation theorems differ in the specific framework they use and the exact utility representation of preferences they yield. The crux of the matter, however, is to represent incomplete preferences not by a single utility with values in $\mathbb{R}$ but by several utilities, each with values in $\mathbb{R}$. What is known now as a multi-utility (or multi-self) representation of preferences. The general idea of these models is that x is preferred to y if all the utilities are higher for x than for y.

$$x \succ y \Leftrightarrow u(x) > u(y), \forall u \in \mathcal{U}$$

Where $\mathcal{U}$ is the set of all utilities. $\mathcal{U}$ generally has some properties, like compactness. When the utility comparisons go in opposite directions, alternatives are incomparable. The decision maker has incomplete preferences. Nishimura and Ok (2016) gave the most general model. Reflexivity is the only hypothesis needed for their preference representation. It is one of the most complex one in practice. To the best of my knowledge, no one has proposed testable implications of this model.

D.2 Experiments on Incomplete Preferences

Danan and Ziegelmeyer (2006) study indecision in a risky environment. In order to do so, they assume that indecision is akin to a preference for flexibility. If the decision maker does not know what to choose, she will prefer to postpone her choice. More importantly, it is the only case where postponing is valuable, that is when she is ready to pay for flexibility. They implement an experiment for decision under risk with this idea, using lotteries and a bracketing procedure. Subjects could postpone their choice at a cost. All of them had to come back the week after to choose between the alternatives they were undecided. Their design allows them to build an index of incompleteness, roughly based on the number of postponed choices. 59% of subjects have incomplete preferences in their experiments.

Cettolin (2016) implement an experiment testing the completeness axiom with ambiguous prospects. They also test the possibility that intransitive indifference explains their results. As they have noted, they do not elicit incomplete preferences directly, but rather carefully designed their experiments and their choice objects to rule out other models. They perform four different experiments to shore up their result. In the first experiment, subjects have to choose between a risky and an uncertain prospect or to delegate the choice to a random device. In the second experiment, they elicit the preferences between a certain outcome and a risky prospect, to elicit possible prospect theory preferences. In a third experiment, they increase the payoff of the ambiguous events compared to the risky ones. Finally, they modify the first experiment by replacing the uncertain prospect by a certain amount. Around half of the subject chose to delegate to the random device, indicating that a significant proportion of decision makers might have incomplete preferences in the risky and uncertain domain. Almost no subject delegate with a certain amount, ruling out intransitive indifference as an explanation for their results.

Sautua (2017) also implements a test of Bewley (2002) preferences and contrast it with reference dependent subjective expected utility of Sugden (2003). He allocated different lottery tickets to participants in an experiment and allowed them to switch for a bonus. In classical theories, subjects should switch if the expected gains are the same. Many subjects switch, however, indicating some degree of incompleteness and some degree of reference-dependent. Again this experiment is concerned with incompleteness in the ambiguous domain.

Qiu and Ong (2017) implement an experiment around the ultimatum game. Proposers face binary choices between the equal allocation and (random) series of unequal allocation. They are allowed to randomize these propositions. Finally, they can state a willingness to pay to get the randomized choice. Receivers face binary choices between accepting and rejecting the proposed allocations. They can randomize between accepting and rejecting. Finally, they can state a willingness to pay to get the randomized choice. Randomizing is only understandable if decision makers have incomplete preferences, as is a positive WTP. They found that 92% of subjects randomize at some point, a strong indication that people have incomplete preferences.

Costa-Gomes et al. (2016) study incomplete preferences through choice deferral, as in Danan and Ziegelmeyer (2006), but based on a model of Gerasimou (2017). They compare four models of decision making, including traditional preference maximization and rational indecision.

Using a modified Houtman-Maks index, they identify from which preference order each observed choices are and therefore to which model decision makers are the closest. 42.7% of the data is closest to traditional utility maximization. The other models explain 21% to rational indecision and the rest. They were 28% of utility maximizers and 5% of rationally indecisive decision makers.

Their experiment is closed to this one, as subjects chose in the powerset of five alternatives (different headphones close in values). They were two treatments, one with forced choice (choice function) and one with the possibility to defer choice (choice correspondence). Deferring was costly, and a significant proportion of decision makers postponed their choice. Deferring offered the possibility to look at each headphone and test them. One question raised in this experiment is whether they have investigated incomplete preferences or intransitive indifference. Indeed, the alternatives are similar to each other, and the main reason for postponing choice might be that they are not distinguishable by the decision maker during the first choice.

Like these authors, I explore incomplete preferences eliciting a choice correspondence instead of a choice function. Unlike them, I do so by directly asking subjects for an ε-correspondence, and in some cases, I identify the choice correspondence of decision makers. Like Costa-Gomes et al. (2016), I investigate incomplete preferences in the certain domain. Unlike most of them, choosing more than one alternative is not costly, on the contrary, it has a positive impact on the gains. Importantly, compared to the risky and uncertain domain experiments, pay-for-certainty does not rely on a precise choice of the choice objects.

D.3 Partial Order Rationalizability

Aleskerov, Bouyssou, and Monjardet (2007) characterize partial orders rationalizability by three axioms: outcast[29], concordance[30], and heredity[31].

[29] also known as Chernoff (1954)'s postulate 5.

[30] The axiom has various names in the literature: γ in Sen (1971), concordance in Aleskerov, Bouyssou, and Monjardet (2007), postulate 5 in Chernoff (1954), and expansion in Moulin (1985).

[31] Heredity is also known as Sen (1971)'s α, Chernoff (1954)'s postulate 4, Arrow (1959)'s condition 3.

Axiom D.1. (α) If an alternative is chosen in a set, it is also chosen in any subset.

$$\text{for all } S, T \in 2^X, (x \in T \subseteq S \text{ and } x \in c(S)) \Rightarrow x \in c(T)$$

Axiom D.2 (Concordance)**.** A choice correspondence satisfies concordance, i.e., γ, if and only if any alternative chosen in two different sets is chosen in the unions:

$$\text{for all } S, S' \in 2^X, c(S) \cap c(S') \subseteq c(S \cup S')$$

γ is a bottom-up condition: choices in small sets constrain choices in large sets.

Axiom D.3 (Outcast)**.** A choice correspondence satisfies outcast if and only any subset S' of S is made of alternatives that are not chosen in S, then removing these alternatives from S does not change the chosen alternatives.

$$\text{for all } S, S' \in 2^X, c(S) \subseteq S' \subseteq S \Rightarrow c(S') = c(S)$$

Or equivalently:

$$\text{for all } S, S' \in 2^X, S' \subseteq S \backslash c(S) \Rightarrow c(S \backslash S') = c(S)$$

Outcast is akin to the independence of irrelevant alternatives.

Aleskerov, Bouyssou, and Monjardet (2007)'s theorem 2.9 shows that a choice correspondence jointly satisfies outcast, heredity (α) and concordance (γ) if and only if a partial order rationalizes it.

References

Agranov, Marina and Pietro Ortoleva (2017). “Stochastic Choice and Preferences for Randomization”. In: *Journal of Political Economy* 125.1, pp. 40–68. ISSN: 0022-3808. DOI: 10.1086/689774.

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.

Andreoni, James and Charles Sprenger (2012). “Estimating Time Preferences from Convex Budgets”. In: *American Economic Review* 102.7, pp. 3333–3356. ISSN: 0002-8282. DOI: 10.1257/aer.102.7.3333.

Anscombe, F. J. and R. J. Aumann (1963). “A Definition of Subjective Probability”. In: *The Annals of Mathematical Statistics* 34.1, pp. 199–205. ISSN: 00034851. DOI: 10.2307/2991295.

Armstrong, W. E. (1939). “The Determinateness of the Utility Function”. In: *The Economic Journal* 49.195, pp. 453–467. ISSN: 00130133, 14680297. URL: http://www.jstor.org/stable/2224802.

Arrow, Kenneth J. (1959). “Rational Choice Functions and Orderings”. In: *Economica* 26.102, p. 121. ISSN: 00130427. DOI: 10.2307/2550390.

Aumann (1962). “Utility Theory without the Completeness Axiom”. In: *Econometrica* 30.3, pp. 445–462. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1909888.

Bewley, Truman F. (2002). “Knightian decision theory. Part I”. In: *Decisions in Economics and Finance* 25.2, pp. 79–110. ISSN: 1129-6569. DOI: 10.1007/s102030200006.

Bouacida, Elias and Daniel Martin (2017). “Predictive Power in Behavioral Welfare Economics”. working paper or preprint. URL: https://halshs.archives-ouvertes.fr/halshs-01489252.

Camerer, Colin F. and Teck-Hua Ho (1994). “Violations of the betweenness axiom and nonlinearity in probability”. In: *Journal of Risk and Uncertainty* 8.2, pp. 167–196. ISSN: 0895-5646. DOI: 10.1007/BF01065371.

Caplin, Andrew, Mark Dean, and Daniel Martin (2011). “Search and Satisficing”. In: *American Economic Review* 101.7, pp. 2899–2922. DOI: 10.1257/aer.101.7.2899.

Cettolin, Elena (2016). “Revealed Incomplete Preferences Under Uncertainty: Evidence for Bewley Preferences”. In: *SSRN Electronic Journal* 11429. ISSN: 1556-5068. DOI: 10.2139/ssrn.2816025.

Chernoff, Herman (1954). “Rational Selection of Decision Functions”. In: *Econometrica* 22.4, p. 422. ISSN: 00129682. DOI: 10.2307/1907435.

Chew, S. H. (1989). “Axiomatic utility theories with the betweenness property”. In: *Annals of Operations Research* 19.1, pp. 273–298. ISSN: 1572-9338. DOI: 10.1007/BF02283525.

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 et al. (2016). "Choice, Deferral and Consistency". In:

Danan, Eric (2008). "Revealed preference and indifferent selection". en. In: *Mathematical Social Sciences* 55.1, pp. 24–37. ISSN: 01654896. DOI: 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.

Danan, Eric and Anthony Ziegelmeyer (2006). *Are Preferences Complete?: An Experimental Measurement of Indecisiveness Under Risk*. URL: https://78462f86-a-4154f0ba-s-sites.googlegroups.com/a/ericdanan.net/www/research/preferences.pdf.

Dekel, Eddie (1986). "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom". In: *Journal of Economic Theory* 40.2, pp. 304–318. ISSN: 00220531. DOI: 10.1016/0022-0531(86)90076-1.

Deparis, Stéphane (2013). "The effect of multicriteria conflict on the expression of preferences : an empirical study". PhD thesis. École Centrale Paris.

Dhar, Ravi and Itamar Simonson (2003). "The Effect of Forced Choice on Choice". In: *Journal of Marketing Research* 40.2, pp. 146–160. ISSN: 0022-2437. DOI: 10.1509/jmkr.40.2.146.19229.

Dohmen, Thomas et al. (2011). "Individual risk attitudes: Measurement, determinants, and behavioral consequences". In: *Journal of the European Economic Association* 9.3, pp. 522–550. ISSN: 15424766. DOI: 10.1111/j.1542-4774.2011.01015.x.

Eliaz, Kfir and Efe A. Ok (2006). "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences". en. In: *Games and Economic Behavior* 56.1, pp. 61–86. ISSN: 08998256. DOI: 10.1016/j.geb.2005.06.007.

Fischbacher, Urs (2007). "z-Tree: Zurich toolbox for ready-made economic experiments". In: *Experimental Economics* 10.2, pp. 171–178. ISSN: 1386-4157. 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. ISSN: 0030-364X. DOI: 10.1287/opre.18.2.207.

Frick, Mira (2016). "Monotone threshold representations". In: *Theoretical Economics* 11.3, pp. 757–772. ISSN: 19336837. DOI: 10.3982/TE1547.

Gerasimou, Georgios (2017). "Indecisiveness, Undesirability and Overload Revealed Through Rational Choice Deferral". In: *The Economic Journal* 1988, pp. 1–30. ISSN: 00130133. DOI: 10.1111/ecoj.12500.

Greiner, Ben (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". In: *Journal of the Economic Science Association* 1.1, pp. 114–125. ISSN: 2199-6776. DOI: 10.1007/s40881-015-0004-4.

Laslier, Jean-François and Remzi Sanver (2010). *Handbook on Approval Voting*. Ed. by Jean-François Laslier and Remzi Sanver. Vol. 1. Studies in Choice and Welfare. Berlin, Heidelberg: Springer. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004.

Luce, R. Duncan (1956). "Semiorders and a Theory of Utility Discrimination". In: *Econometrica* 24.2, pp. 178–191. ISSN: 0012-9682. 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: 10.1016/j.geb.2004.02.007.

Manzini, Paola and Marco Mariotti (2012). "Categorize Then Choose: Boundedly Rational Choice and Welfare". en. In: *Journal of the European Economic Association* 10.5, pp. 1141–1165. ISSN: 1542-4774. DOI: 10.1111/j.1542-4774.2012.01078.x.

Moulin, H. (1985). "Choice functions over a finite set: A summary". In: *Social Choice and Welfare* 2.2, pp. 147–160. ISSN: 0176-1714. DOI: 10.1007/BF00437315.

Nehring, Klaus (1997). "Rational choice and revealed preference without binariness". In: *Social Choice and Welfare* 14.3, pp. 403–425. ISSN: 0176-1714. DOI: 10.1007/s003550050075.

Nishimura, Hiroki and Efe A. Ok (2016). "Utility representation of an incomplete and nontransitive preference relation". In: ISSN: 00220531. DOI: 10.1016/j.jet.2016.07.002.

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: 10.1006/jeth.2001.2814.

Ok, Efe A., Pietro Ortoleva, and Gil Riella (2015). "Revealed (P)Reference Theory". en. In: *American Economic Review* 105.1, pp. 299–321. ISSN: 0002-8282. DOI: 10.1257/aer.20111046.

Qiu, Jianying and Qiyan Ong (2017). "Indifference or Indecisiveness: a strict Discrimination". In: 8225. ISSN: 0148-396X. DOI: 10.1227/01.NEU.0000349921.14519.2A.

Samuelson, Paul A (1938). "A Note on the Pure Theory of Consumer's Behaviour". In: *Economica* 5.17, p. 61. ISSN: 00130427. 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: 01672681. DOI: 10.1016/j.jebo.2017.02.003.

Savage, Leonard J (1954). *The foundations of statistics*. Courier Corporation.

Sen, Amartya Kumar (1971). "Choice Functions and Revealed Preference". In: *The Review of Economic Studies* 38.3, pp. 307–317. ISSN: 0034-6527. DOI: 10.2307/2296384.

— (1997). "Maximization and the Act of Choice". In: *Econometrica* 65.4, p. 745. ISSN: 00129682. DOI: 10.2307/2171939.

Sugden, Robert (2003). "Reference-dependent subjective expected utility". In: *Journal of Economic Theory* 111.2, pp. 172–191. ISSN: 00220531. DOI: 10.1016/S0022-0531(03)00082-6.

Tyson, Christopher J. (2018). "Rationalizability of menu preferences". In: *Economic Theory* 65.4, pp. 917–934. ISSN: 0938-2259. DOI: 10.1007/s00199-017-1043-2.

Von Neumann, John and Oskar Morgenstern (1953). *Theory of Games and Economic Behavior: 3d Ed*. Princeton University Press.