DISCUSSION PAPER SERIES

IZA

IZA DP No. 2341

Two-Stage Boundedly Rational Choice Procedures: Theory and Experimental Evidence

Paola Manzini
Marco Mariotti

September 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Two-Stage Boundedly Rational Choice Procedures: Theory and Experimental Evidence

Paola Manzini
Queen Mary, University of London and IZA Bonn

Marco Mariotti
Queen Mary, University of London

Discussion Paper No. 2341
September 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2341
September 2006

ABSTRACT

Two-Stage Boundedly Rational Choice Procedures: Theory and Experimental Evidence*

We study and test a class of boundedly rational models of decision making which rely on sequential eliminative heuristics. We formalize two sequential decision procedures, both inspired by plausible models popular among several psychologists and marketing scientists. However we follow a standard 'revealed preference' economic approach by fully characterizing these procedures by few, simple and testable conditions on observed choice. Then we test the models (as well as the standard utility maximization model) with experimental data. We find that the large majority of individuals behave in a way consistent with one of our procedures, and inconsistent with the utility maximization model.

JEL Classification: C91, D9

Keywords: bounded rationality, choice experiments

Corresponding author:

Marco Mariotti
Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
E-mail: m.mariotti@qmul.ac.uk

* Part of this work was carried out while Manzini and Mariotti were visiting the University of Trento, and funding for the experiments was provided by the ESRC under grant RES-000-22-0866. We wish to thank both institutions for their support. We are also grateful to Michele Lombardi, Luigi Mittone, Daniel Read and Nick Vriend for insightful discussions and comments, as well as to the tireless staff of the CEEL lab in Trento, in particular to Marco Tecilla for superb programming support. All errors are our own.

1 Introduction

In the standard model of decision making choice behavior is described as the outcome of the maximization of some binary relation, possibly summarized by a utility function. Yet, observed choice behavior is often incompatible with this model.[1] In this paper we study and test a class of boundedly rational models of decision making which rely on *sequential eliminative heuristics*. We formalize two sequential decision procedures, both inspired by plausible models popular among several psychologists and marketing scientists. However we follow a standard 'revealed preference' economic approach by fully characterizing these procedures by few, simple and testable conditions on observed choice. Then we test the models (as well as the standard utility maximization model) with experimental data. We find that the large majority of individuals behave in a way consistent with one of our procedures, and inconsistent with the utility maximization model. Our theoretical results also allow us to trace the observed departures from maximization to either of two elementary forms of inconsistency, and therefore to guide the search for any alternative model.

For the moment let us refer to our procedures as Procedure I and Procedure II.[2] For an informal example of these procedures, suppose a decision-maker (DM) has to choose which of three wines to drink with his dinner at a restaurant: a moderately priced Australian (A), an expensive French Bordeaux (B), or a cheap Italian Chianti (C). Procedure I describes considerations of the following kind. DM looks first at the origin: he thinks B is best among European wines, so he prefers B to C, but has never tasted new world wines, so cannot compare A to either B or C. When he is undecided, DM chooses according to price, and C is cheaper than A which is cheaper than B. This generates the following 'irrational' behavior: if only A and B are available DM chooses A; if only B and C are available he chooses B; finally if only A and C are available he chooses C. DM exhibits *pairwise cyclical* choices.

Procedure II captures considerations of the following kind. When all wines are avail-

[1]See e.g. Roelofsma and Read [26], Tversky [35], and Waite [38] who find evidence of pairwise cycles of choice. The evidence presented in this paper points to further violations of rational behavior.

[2]Procedure I was first introduced in Manzini and Mariotti [20].

able, the DM perceives B and C as a group of 'similar' wines, the 'old world' wines. When comparing groups, DM prefers the 'new world' Australian (degenerate) group to the 'old world' group. However when only binary choices are available, DM looks straight at price. This generates the following additional type of irrational behavior: A is selected when all three wines are available, but C is selected (based on price) in both binary contests in which it appears. DM exhibits a strong form of *menu dependence*.

We shall see below that the two types of 'rationality failure' just illustrated are an exhaustive taxonomy also for general choices.

Formally, in both procedures, DM implements a *two-stage algorithm* to arrive at a final choice: two *rationales* (asymmetric and possibly incomplete, binary relations) are used sequentially to eliminate alternatives from the choice set. The difference lies in the way the first rationale operates. In procedure I, the first rationale simply ranks (some of) the alternatives. In procedure II, the first rationale ranks instead *sets* of alternatives. The interpretation is that the decision-maker perceives some alternatives as similar in some aspect and treats them as a group (such as the 'old world' wines in the example). In general there may be more than one similarity aspect, so groups might overlap. When two similarity groups are disjoint, they may be related by the first rationale. In this case, the entire 'losing' group is eliminated. For both procedures, the second rationale is used to eliminate further alternatives and to single out a choice.

Surprisingly, Procedure II is fully characterized by a single 'revealed preference' property, which we call WARP* (Theorem 9). WARP* is a simple weakening of the standard Weak Axiom of Revealed Preference (Samuelson [30]). WARP* adds to WARP the clauses between brackets in the following definition: if x is directly revealed preferred to y [both in pairwise contests and in the presence of a 'menu' of other alternatives], then y cannot be directly revealed preferred to x [in the presence of a smaller menu]. This characterization is one of the main contributions of our work. It shows how the analysis of non-trivial forms of bounded rationality is amenable to tests on observed behavior of exactly the same kind, and roughly as simple, as the tests used to check full rationality (maximization of a transitive and complete binary relation).

Procedure I can be similarly characterized, as we have shown in Manzini and Mariotti

[20]. The choice data can be generated by procedure I if and only if they satisfy, in addition to WARP*, another standard revealed preference property called Expansion. Expansion says that if x is chosen from two sets of alternatives, then x is chosen when the sets are merged. Expansion and WARP* together still constitute a weakening of WARP, which implies both properties.

The nature of these procedures can be much better understood in the light of a simple but crucial result on choice functions. Any failure of full rationality in choice (that is, a violation of WARP) can be ultimately reduced to one (or both) of just two categories, illustrated in the previous wine examples: pairwise inconsistent choice and 'Condorcet inconsistent' choice. Condorcet inconsistency captures a specific type of 'menu-effect' in choice: an alternative is chosen over each of a number of other alternatives in pairwise choice, yet it is no longer chosen when all these alternatives are grouped together.[3] On the other hand, pairwise inconsistent choice involves only pairwise comparisons and therefore does not incorporate any menu-effect. It thus captures a conceptually separate violation of full rationality, namely that the pairwise comparisons do not allow the observer to construct a preference relation with some minimal consistency property. Procedure I can address violations of pairwise cyclical choice, but not of Condorcet inconsistency (which is a necessary condition for choices generated by that procedure). Procedure II can explain both types of irrational behavior. Needless to say, even Procedure II is not vacuous as it can be tested by WARP*.

After developing our theoretical analysis, we put all the decision models discussed to the experimental test. To do so we elicit the choice function of experimental subjects out of all possible subsets of a given initial set of alternatives. In many decision theory experiments only pairwise choices (or preferences) are elicited. However, the taxonomy result discussed above illustrates the importance of observing decision behavior on larger sets in order to trace the sources of full rationality violations.

In our experiment we use as alternatives time sequences of monetary rewards. Our

[3]Obviously there are other specifications of what constitutes menu-dependence, which for example could be 'dynamic' (dependence on previous choices or status quo: see e.g. Masatlioglu and Ok [22], Houy [15] and Botond and Koszegi [17]). See Sen [32] for a general discussion of menu-dependence.

data show that *WARP is violated by a majority of subjects.* One of our main contributions on the experimental front is an enquiry into the specific nature of the violations of full rationality, based on our taxonomy result. In our context, more than 15% pairwise cyclical choices were observed. Nonetheless these violations of full rationality were strongly associated with violations of Expansion. The consequence of this fact is that Procedure I does not make a significant improvement of the standard maximization model (of course, Procedure I may be more useful for other purposes[4]).

Our main experimental finding is that Procedure II yields a step change in explanatory power in the present case. The majority of violations of WARP are associated with Condorcet inconsistency, so that any successful model will need to incorporate this type of menu effect. Indeed, the *large majority of subjects satisfies WARP** in all their choices, thus validating Procedure II as a boundedly rational model of decision making.

2 Theory

2.1 Preliminaries

Let X be a finite set of alternatives and let $\Sigma \subseteq 2^X$. A choice function on Σ is a function $\gamma : \Sigma \rightarrow X$, such that $\gamma(S) \in S$ for all $S \in \Sigma$. The only additional assumptions[5] we make on the domain Σ are that, for all $x, y, z \in X$:

(i) $\{x, y\} \in \Sigma$;

(ii) $\{x, y, z\} \in \Sigma$.

For a binary relation $B \subseteq X \times X$ denote the $B-$maximal elements of a set $S \in \Sigma$ by $\max(S, B)$, that is:

$$\max(S, B) = \{x \in S | \nexists y \in S \text{ for which } (y, x) \in B\}$$

[4]In Manzini and Mariotti [21] we study Procedure I with *specific* decision criteria and obtain a model of time preferences with significantly higher explanatory power than discounting models. In general, it is a consequence of our theoretical results that when only *pairwise* choices are the object of interest, Procedure I is as a good a model as Procedure II.

[5]The finiteness assumption on the domain could be easily dispensed with, and replaced by a well-behavedness assumption on the sets in Σ, guaranteeing that complete and transitive relations on them have a maximal element. This would just complicate notation so we stick with the finite case in the text.

Definition 1 *A choice function is fully rational if there exists a complete order*[6] $B \subseteq X \times X$ *such that* $\gamma(S) = \max(S, B)$ *for all* $S \in \Sigma$.

As is well-known[7], in the present context the fully rational choice functions are exactly those that satisfy WARP, defined below:

WARP: If $x = \gamma(S)$, $y \in S$ and $x \in T$ then $y \neq \gamma(T)$.

WARP says that if an alternative is directly revealed preferred to another, the latter alternative can never be directly revealed preferred to the former (revealed preference is an asymmetric relation).

2.2 A taxonomy of irrationality

Failures of full rationality may mix together more than one elementary form of inconsistency. To reduce lack of full rationality to its basic building blocks (to be studied in the experiment), we consider the two violations of WARP discussed in the introduction, menu dependence and pairwise inconsistency. The latter category involves exclusively choices between *pairs* of alternatives. The former category instead involves choices from larger sets.

The following property captures the elementary form of *menu independence*:

Condorcet consistency: If $x = \gamma(\{x, z\})$ for all $z \in S \setminus \{x\}$ and $S \in \Sigma$, then $x = \gamma(S)$.

Condorcet consistency says that if the same alternative is chosen in pairwise contests against any other alternative in a set, then this alternative will be chosen from the set.

Let P_γ denote the *base relation* of a choice function γ, that is $(x, y) \in P_\gamma$ if and only if $x = \gamma(\{x, y\})$. Each of the following properties capture elementary forms of pairwise consistency:

Base transitivity: P_γ is transitive.

---

[6]That is, a transitive binary relation.

[7]See e.g. Moulin [24] or Suzumura [34]

Base acyclicity: P_γ is acyclic.

Base intervality: P_γ satisfies the intervality condition:[8] If $(x,y),(w,z) \in P_\gamma$ then either $(x,z) \in P_\gamma$ or $(w,y) \in P_\gamma$

We note first that all these conditions are equivalent:

Proposition 2 *A choice function satisfies base intervality, if and only if it satisfies base transitivity, if and only if it satisfies base acyclicity.*

Proof: Base acyclicity $\Rightarrow$ Base intervality. Suppose that γ violates base intervality, that is there exists $x,y,w,z \in X$ for which $(x,y),(w,z) \in P_\gamma$ but $(x,z),(w,y) \notin P_\gamma$. Suppose first that $x \neq z$ and $w \neq y$. Then since $\{x,z\},\{w,y\} \in \Sigma$ it must be $(z,x),(y,w) \in P_\gamma$. Therefore we have constructed the base cycle $(x,y),(y,w),(w,z),(z,x) \in P_\gamma$. Suppose next that $x = z$. Then we have the cycle $(x,y),(y,w),(w,x) \in P_\gamma$. Similarly for the case $y = z$.

Base intervality $\Rightarrow$ Base transitivity. Suppose that γ violates base transitivity, so that there exists $x,y,z \in X$ for which $(x,y),(y,z) \in P_\gamma$ but $(x,z) \notin P_\gamma$. By the single-valuedness of γ it cannot be $x = z$, and by the fact that $\{x,z\} \in \Sigma$ it must be $(z,x) \in P_\gamma$. Now we have $(x,y),(z,x) \in P_\gamma$ but also $(x,x) \notin P_\gamma$ and $(z,y) \notin P_\gamma$, violating base intervality.

That base transitivity implies base acyclicity is obvious. ■

The equivalence of all the base conditions makes it legitimate to speak simply of *Pairwise Consistency* when the three conditions are met. The following is our basic classification result:

Theorem 3 *A choice function satisfies WARP if and only if satisfies both Condorcet consistency and Pairwise Consistency.*

Proof: It is obvious that a choice function that violates Condorcet consistency also violates WARP. Suppose it violates base transitivity. Then by the domain assumption (i)

[8] Introduced in Fisburn [9].

there exists $x, y, z \in X$ for which $(x, y), (y, z), (z, x) \in P_\gamma$, and by the domain assumption (ii) WARP is contradicted on $\{x, y, z\}$.

For the converse implication, suppose that γ violates WARP, and let $S, T \in \Sigma$ be such that $x = \gamma(S) \neq y = \gamma(T)$, $x, y \in T \cap S$. Suppose that γ satisfies base transitivity: we show that then γ must violate Condorcet consistency. By base transitivity there exist base-maximal elements in S and T, that is $s \in S$ and $t \in T$ such that

$$s = \gamma(\{s, z\}) \text{ for all } z \in S \backslash \{s\}$$

$$t = \gamma(\{t, z\}) \text{ for all } z \in T \backslash \{t\}$$

If $s \neq x$, then Condorcet consistency is violated on S. If $s = x$, then in particular $x = \gamma(\{x, y\})$. So $y \neq t$ and Condorcet consistency is violated on T. ■

2.3 Two sequential models of decision making

We first explain our Procedure I, introduced in Manzini and Mariotti [20]. From now on we shall call any asymmetric binary relation on $X \times X$ a *rationale.*

Definition 4 *A choice function γ is a* **Rational Shortlist Method (RSM)** *if and only if there exists an ordered pair (B_1, B_2) of rationales such that:*

$$\{\gamma(S)\} = \max(\max(S, B_1), B_2) \text{ for all } S \in \Sigma$$

In that case we say that (B_1, B_2) sequentially rationalize γ.

So the choice from each S can be represented as if the decision maker went through two sequential rounds of elimination of alternatives. In the first round he retains only the elements which are maximal according to rationale B_1. In the second round, he retains only the element which is maximal according to rationale B_2: that element is his choice. Note that, crucially, the rationales and the sequence are invariant with respect to the choice set. RSM's are in the vein of several 'noncompensatory' sequential eliminative heuristics[9]

[9] The adjective 'noncompensatory' refers to the fact that while several criteria are used, they cannot be used to 'compensate' each other: unlike the arguments of a utility function, there is no trade-off in the mind of the decision maker between one criterion and the next.

promoted by psychologists (such as the Elimination by Aspects model by Tversky [36], or the 'Fast and frugal heuristics' by Gigerenzer and Todd [11]) and marketing scientist (such as 'greedoid based' choice algorithms, Yee, Dahan, Hauser and Orlin [39]*)*.

RSM are characterized by two axioms, one which is a weakening of WARP, and the other is classical Expansion axiom:

WARP*: For all $R, S \in \Sigma$: If $\{x, y\} \subset R \subset S$ and $x = \gamma(\{x, y\}) = \gamma(S)$ then $y \neq \gamma(R)$.

Expansion: For all $R, S \in \Sigma$ with $S \cup T \in \Sigma$: If $x = \gamma(S) = \gamma(T)$ then $x = \gamma(S \cup T)$.

The following characterization result is an easy corollary of Theorem 2 in Manzini and Mariotti [20].

Theorem 5 *Suppose the domain Σ is closed under set union. Then a choice function on Σ is an RSM if and only if it satisfies WARP* and Expansion.*

Next, we consider 'Procedure II' from the Introduction.

Definition 6 *A* **rationale by similarities** *on X is an asymmetric relation $B \subseteq 2^X \times 2^X$ satisfying the following two properties:*

(i) $R \cap S = \varnothing$ whenever $(R, S) \in B$

(ii) $|R \cup S| > 2$ whenever $(R, S) \in B$

The interpretation of a rationale by similarities B is that some alternatives are grouped by similarity in some aspect. Similarity can be in more than one aspect, hence two similarity groups are not necessarily disjoint. However, two conditions suggest the similarity interpretation. First, the decision-maker can only compare disjoint groups (condition (i)). Moreover, what the relation compares are genuinely groups: degenerate comparisons between singletons are not allowed (condition (ii)). For example, constant streams of monetary payments can form a group 'against' increasing streams, and increasing three-period streams could form a group against increasing two-period streams. The characterization result below would hold without these restrictions.[10]

[10]Rubinstein [27] (and more recently [28]) pioneered in economic theory the analysis of similarity considerations in decision-making. In his work, Rubinstein axiomatizes directly a similarity relation.

Definition 7 *Given a rationale by similarity B and $S \in \Sigma$, the $B-$maximal set on S is given by:*

$$\max(S, B) = \{x \in S| \text{ for no } R', R'' \subseteq S \text{ it is the case that } (R', R'') \in B \text{ and } x \in R''\}$$

We are now ready for our main definition:

Definition 8 *A choice function γ is* **two-rational by similarities** *if and only if there exists a rationale by similarities B_1 and a rationale B_2 such that:*

$$\{\gamma(S)\} = \max(\max(S, B_1), B_2) \text{ for all } S \in \Sigma$$

So, the decision maker looks first at group rankings, and eliminates any group which is dominated by another group. Then, she decides among the remaining alternatives on the basis of the second rationale. For example, if the choice set is comprised of two decreasing streams of money and two increasing streams, the decision maker may first select the group of decreasing streams and then select within that group. When this procedure leads to a single chosen alternative for each choice set, the resulting choice function is two-rational by similarities.

Below we make some observations which highlight key differences between the two sequential procedures we have introduced.

Remark 1 *Condorcet consistency is a necessary condition for an RSM. However, there are choice functions which violate Condorcet consistency and yet are two-rational by similarities. For instance, take the following choice function, with the base relation as indicated in figure 1: $X = \{a, b, c, d\}$, $\gamma(X) = \gamma(\{a, b, c\}) = \gamma(\{b, c, d\}) = b$, $\gamma(\{a, c, d\}) = \gamma(\{a, b, d\}) = a$. Condorcet consistency is violated, since a is chosen in pairwise comparisons over each of the other alternatives but is not chosen from the grand set, nor from $\{a, b, c\}$. However, this choice function is two-rational by similarities with $B_1 = \{(\{b\}, \{a, c\})\}$, and B_2 coinciding with the base relation.*

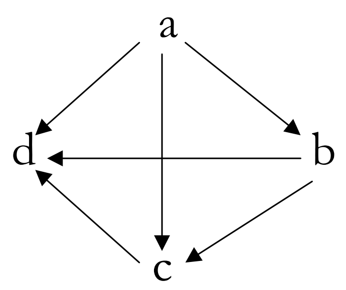

Figure 1: Base relation for remarks 1 and 2

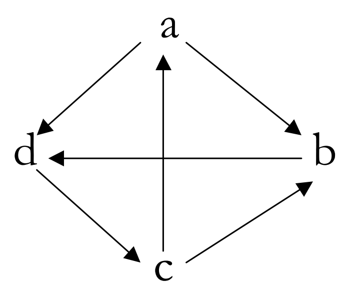

Figure 2: Base relation for remark 3

Remark 2 *Consider the following dual property to Condorcet consistency: an alternative is not chosen in a set in which it is never chosen in any pairwise choice. Formally, if* $x \neq \gamma(\{x,z\})$ *for all* $z \in S \backslash \{x\}$ *and* $S \in \Sigma$, *then* $x \neq \gamma(S)$. *It is easy to show that this property too is a necessary condition for an RSM. However, there exist choice functions that violate it and yet are two-rational by similarities. Consider the choice function with the same base relation as the one in figure 1, where* $\gamma(X) = \gamma(\{a,b,c\}) = \gamma(\{b,c,d\}) = b$, $\gamma(\{a,b,d\}) = a$ *and* $\gamma(\{a,c,d\}) = d$. *This choice function violates the property and is two-rational by similarities by the two asymmetric relations* $B_1 = \{(\{b\},\{a,c\}),(\{d\},\{a,c\})\}$ *and* B_2 *coinciding with the base relation.*

Remark 3 *Two-rationality by similarities is not a vacuous notion: there are choice functions which are not two-rational by similarities. As an example, let* $X = \{a,b,c,d\}$ *and let* $\gamma(X) = \gamma(\{a,b,c\}) = a$, $\gamma(\{a,b,d\}) = \gamma(\{b,c,d\}) = b$ *and* $\gamma(\{a,c,d\}) = c$, *with the base relation as in figure 2. Then, since* b *is chosen in* $\{a,b,d\}$, *there must be* $(R', R'') \in B_1$, *with* $a \in R''$, *so that* a *is eliminated before it can eliminate* b. *Then two-rationality by similarities is made impossible by* $a = \gamma(X)$.

The choice function in the last example violates WARP*. As we show next, this property alone characterizes two-rationality by similarities. To ease notation, fixing the choice function γ, we define the upper and lower contour sets of an alternative on a set $S \in \Sigma$ as

$$Up(x,S) = \{y \in X | \, (y,x) \in P_\gamma\} \cap S$$

and

$$Lo(x,S) = \{y \in X | \, (x,y) \in P_\gamma\} \cap S$$

respectively.

Theorem 9 *A choice function is two-rational by similarities if and only if it satisfies WARP*.*

Proof: Necessity. Suppose that γ is two-rational by similarities with rationales B_1 and B_2. Suppose $x = \gamma(\{x, y\})$ and $x = \gamma(S)$ with $y \in S$. Now suppose by contradiction that $y = \gamma(R)$ with $x \in R \subset S$. This means that x must be eliminated in the first round of elimination in R: if not, then either x would eliminate y in the second round, or $\gamma(\{x, y\}) = x$ would contradict the assumption that γ is rationalized by B_1 and B_2. In particular there exist $R', R'' \subseteq R$, such that $(R', R'') \in B_1$ and $x \in R''$. Since $R', R'' \subset S$ this contradicts $x = \gamma(S)$.

Sufficiency. Define: $(x, y) \in B_2$ if and only if $x = \gamma(\{x, y\})$. B_2 is obviously asymmetric. Define: $(R, S) \in B_1$ if and only if there exists $T \in \Sigma$ such that

$$R = \{\gamma(T)\} \cup Lo(\gamma(T), T)$$

and

$$S = Up(\gamma(T), T) \neq \varnothing$$

B_1 is also obviously asymmetric and note that $R \cap S = \varnothing$ whenever R and S are related by B_1.

Now let $S \in \Sigma$ and let $x = \gamma(S)$. We show that x is not eliminated in either round. Suppose first that $(y, x) \in B_2$ for some $y \in S$. Then by construction

$$(\{x\} \cup Lo(x, S), Up(x, S)) \in B_1$$

and y is eliminated in the first round.

Next, suppose by contradiction that x is eliminated in the first round. Then there exists $R', R'' \subset S$ with $(R', R'') \in B_1$ and $x \in R''$. Define $R = R' \cup R''$. By construction of B_1 it must be

$$R' = \{\gamma(R)\} \cup Lo(\gamma(R), R)$$

and

$$R'' = Up(\gamma(R), R)$$

This means that

$$x = \gamma(\{x, \gamma(R)\})$$

Together with $x = \gamma(S)$ (and noting that $R = R' \cup R'' \subseteq S$) this contradicts WARP*. It remains to note that y is eliminated either in the first round or in the second round for all $y \neq x$. If $y = \gamma(\{x, y\})$, then y is eliminated in the first round. If $x = \gamma(\{x, y\})$ then y is eliminated in the second round since as we have seen x survives the first round. ■

3 Experiment

3.1 Experimental Design

The experiment was carried out at the Computable and Experimental Economics Laboratory at the University of Trento, in Italy. We ran a total of 13 sessions between July 2005 and February 2006. Participants were recruited through bulletin board advertising from the student population of the University of Trento. Male and female participants took part in each experimental session in roughly equal proportions. The experiment was computerised, and each participant was seated at an individual computer station, using separators so that subjects could not see the choices made by other participants. Experimental sessions lasted an average of around 26 minutes, of which an average of 18 minutes of effective play, with the shortest session lasting approximately 16 minutes and the longest around 37 minutes. We considered two treatments, one in which subjects received only a 5 Euro showup fee (a total 56 subjects in 4 sessions), and one with payments based on choice, where as we explain more in detail below an additional 48 Euros were made available to each subject (a total of 102 subjects in 9 sessions).[11] We will refer to these two treatments as HYP (for hypothetical) and PAY (for paid), respectively.[12] At the beginning of the experiment subjects read instructions on their monitor, while an experimenter read the instructions aloud to the participants.[13] In each treatment,

[11]The show up fee alone, for an average of less than thirty minute long experimental session, was higher than the hourly pay on campus, which is 8 Euros. At the time of the experiments the exchange rate of the Euro was approximately 1Euro=1.2$=0.7£.

[12]Distinguishing by treatment, sessions lasted an average of about 28 minutes for the PAY treatment, of which an average of just above 19 minutes of effective play; and an average of around 22 minutes for the HYP treatment, of which an average of about 16 minutes of effective play.

[13]See the appendix for the translation of the original instructions (in Italian).

each experimental subject was presented with 23 different screens. Each screen asked the subject to choose the preferred one among a set of alternative remuneration plans in installments to be received staggered over a time horizon of nine months, each consisting of 48 Euros overall. Instructions were the same in both treatments, bar for one sentence, which in the HYP treatment clarified that choices were purely hypothetical, so that the only payment to be received would be the show up fee; whereas for the PAY treatment it was explained that at the end of the experiment one screen would be selected at random, and the preferred plan for that screen would be delivered to the subjects.[14]

Choices were based on two sets - depending on the number of installments - of four plans each, namely an increasing (I), a decreasing (D), a constant (K) and a jump (J) series of payments, over either two or three installments, as shown below. Though in both cases payments extended over nine months, because of the different number of installments we abuse terminology and refer to 'two-*period*' and 'three-*period*' sequences rather than two/three-*installment* sequences:

Two period sequences				
	I2	D2	K2	J2
in three months	16	32	24	8
in nine months	32	16	24	40

Three period sequences				
I3	D3	K3	J3	
8	24	16	8	in three months
16	16	16	8	in six months
24	8	16	32	in nine months

Table 1: the base remuneration plans

Each subject had to choose the preferred plan from each possible subset of plans within each group (making up 11 choices per group). In addition, in a 23rd question subjects were asked to choose between the three period sequences SJ=(8,32,8) and SI=(24,8,16). This was needed to address an issue for a different experiment, which we discuss elsewhere.[15]

Figure 3 displays sample screenshots of the choices subject had to make. The participants made their choice by clicking with their mouse on the button corresponding to the preferred remuneration plan. Once made, each choice had to be confirmed, so as to

[14]The experimental lab has a long tradition, so there was no issue of (mis)trust in receiving delayed payments. At the time of writing all subjects have been paid.

[15]See Manzini, Mariotti and Mittone [21]

minimize the possibility of errors. Both the order in which the twenty-three questions appeared on screen and the position of each option on the screen was randomized.

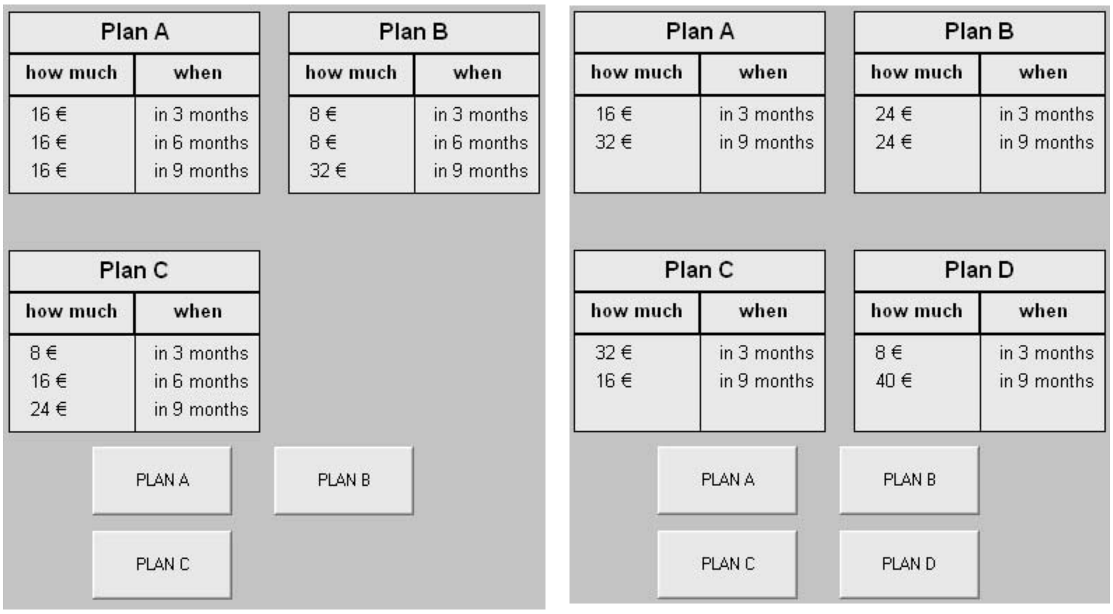

Figure 3: Sample screenshots

In the experiment we elicited the choice functions with domain over all subsets for each of the two grand sets $X2 = \{I2, D2, K2, J2\}$ and $X3 = \{I3, D3, K3, J3\}$. This enables us to check whether or not the axioms discussed in section 2 hold. In particular, we can assess (i) what the main reason is for the failure of full rationality (violation of Pairwise Consistency or violation of Condorcet consistency), and (ii) what proportion of choice functions can be rationalized in the standard way, what proportion is an RSM and what proportion is two-rational by similarities.

3.2 Experimental results I: Evaluating the models

We begin by noting that we can rule out the possibility that experimental subjects choose randomly.[16] Since we are eliciting the entire choice functions from universal sets with four

[16]Purely random choice is an important benchamrk. Within consumer's choice, the idea was first advanced by Becker [4] and it is used for example as the alternative hypothesis in the popular Bronars [5] index of power for nonparametric revealed preference tests. See Andreoni and Harbaugh [2] for a recent

alternatives, with a uniform probability distribution on each choice set, the probability of observing even only two subjects with the same choice is effectively zero for all practical purposes. In fact, as there are a possible $2^6 \cdot 3^4 \cdot 4 = 20,736$ choice functions on each universal set, that probability is $(20,736)^{-2} = 2.325\,7 \times 10^{-9}$. On the contrary for both treatments and for both universal sets $X2$ and $X3$ we find almost half of the subjects with the same modal choice function. For illustration we report the frequency distributions of the observed choice functions only graphically in Figure 4 (we omit labels for legibility).

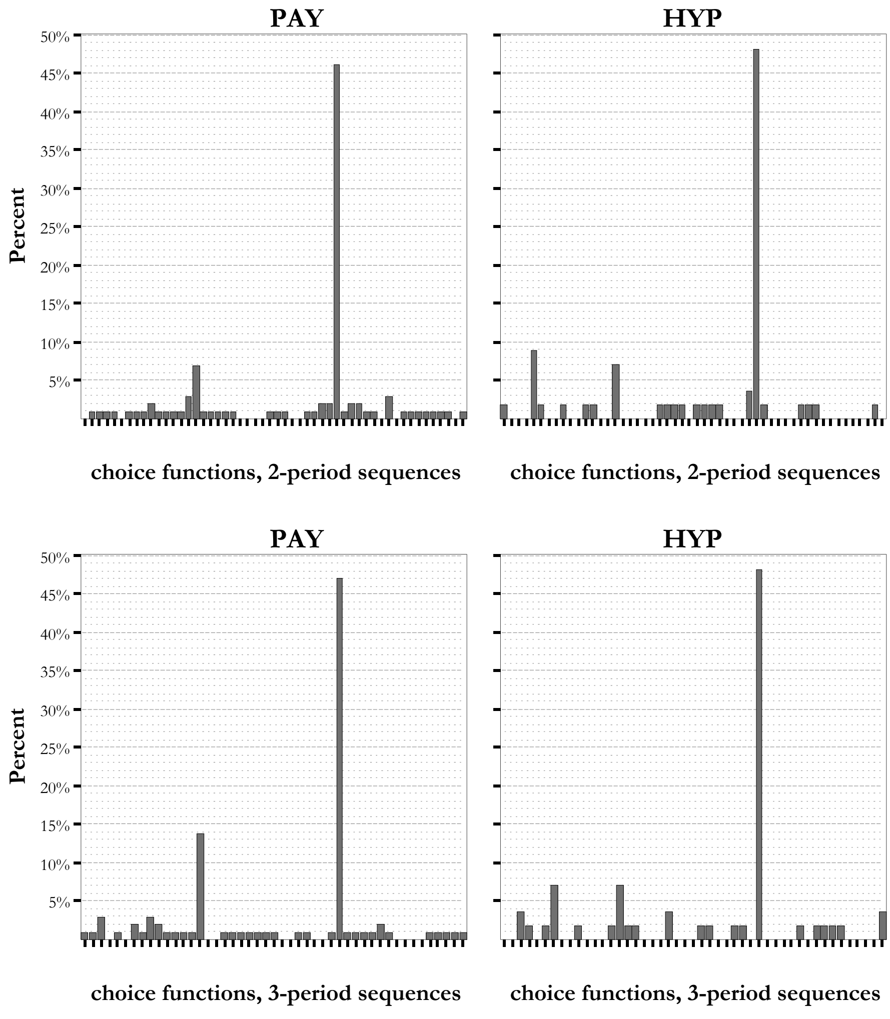

Figure 4: Frequency distributions of choice profiles by treatment and sequence length

discussion of this issue.

The corresponding data are reported in tables 24 and 25 in appendix A.5.

Before turning to the evaluation of the models,[17] we check which, if any, of the two failures of full rationality highlighted in Theorem 3 in section 2.2 is more prevalent. To this effect we begin by reporting aggregate data for the *violations* of Pairwise and Condorcet Consistency for each of the choice functions elicited (i.e. the choice functions from 2^{X2} and from 2^{X3} for each treatment):

	PAY				HYP			
	2 periods		3 periods		2 periods		3 periods	
	#	%	#	%	#	%	#	%
Condorcet Consistency	39	*38.2*	30	*28.4*	14	*25*	12	*21.4*
Pairwise Consistency	12	*11.8*	7	*6.9*	3	*5.4*	3	*5.4*

Table 2: Axiom violations in the two treatments.

Since we are mainly interested in the decisions of individual subjects over *all* their choices, we report below a summary of overall violations of Condorcet Consistency and Pairwise Consistency *by experimental subject*:

	PAY		HYP	
	#	%	#	%
Condorcet Consistency	51	*50*	22	*39.3*
Pairwise Consistency	17	*16.7*	4	*7.1*

Table 3: Overall violations of PC and CC.

From table 3 it emerges that failures of Condorcet Consistency are substantially more frequent than violations of Pairwise Consistency. This difference is statistically significant, regardless of treatment. In fact, the McNemar test of the hypothesis that the proportions of subjects violating Condorcet Consistency is the same as the proportion of subjects violaitng Pairwise Consistency yields exact p-values of 0.009 in the case of the

[17]All the exact statistical analysis has been carried out usting *StatXact*, v.7. For a comprehensive treatment of exact and other methods in categorical data analysis see Agresti [1].

PAY treatment, and of 0.001 in the case of the HYP treatment. If we then look at the differences in the proportion of violations of each of the two axioms *across* treatments, the fall in the proportion of violations when moving from the PAY to the HYP treatment is not statistically significant: Fisher test's exact mid-p values are 0.110 for Condorcet Consistency and 0.470 for Pairwise Consistency.

Next, we turn to the three models examined in section 2.3, and we study the violations of the axioms which characterize those models.[18] Recall that one model is the full rationality model (characterized by WARP), the other is the RSM model (characterized by WARP* and Expansion) and the third is the two-rationality by similarities model (characterized by WARP*). Again we start by looking at data for each choice function, reported in table 4.

	PAY				HYP			
	2 periods		3 periods		2 periods		3 periods	
	#	%	#	%	#	%	#	%
WARP*	22	*21.6*	12	*11.7*	6	*10.7*	5	*8.9*
Expansion	39	*38.2*	30	*29.4*	14	*25*	13	*23.2*
WARP	43	*42.2*	30	*29.4*	16	*28.6*	13	*23.2*

Table 4: Violations of the axioms used for rationalizability.

Table 4 shows that WARP is violated quite considerably in the PAY treatment, and less so - though still substantially - in the HYP treatment. Expansion is violated slightly less often than WARP overall, but much more often than WARP*.

Overall violations by each individual, regardless of choice set, are reported in table 5.

[18] In our experiment we use a small universal set of alternatives. Evidence for choice from budget sets includes Fevrier and Visser [8], Mattei [23] and especially Sippel [33], who find substantial violations of the Generalized Axiom of Revealed Preferences in choices out of budget sets. However, Andreoni and Harbaugh [2] argue that most of these violations are 'small' on the basis of Afriat's efficiency index. Indeed Harbaugh, Krause and Berry [13] and especially Andreoni and Miller [3] find that subjects have choices consistent with GARP in experiments with budget sets.

	PAY		HYP	
	#	%	#	%
Expansion	51	*50*	22	*39.3*
WARP*	29	*28.4*	8	*14.3*
WARP	54	*52.9*	22	*39.3*

Table 5: Overall axiom violations.

The proportion of subjects violating each axiom falls when moving from the PAY to the HYP treatment. Of these differences, those concerning WARP and WARP* are statistically significant, while for Expansion this is not the case (Fisher test's exact mid-p values are 0.110 for Expansion, 0.042 for WARP and 0.022 for WARP*). Like Table 4, Table 5 also suggests similar rates of violation for Expansion and WARP (50% and 52.9%), and substantially lower rates for WARP* compared to either of the other axioms (28.4%). Within treatment, however, the only meaningful comparison in the difference of proportions is between failures of Expansion and WARP*, which are the only two independent axioms.[19] Here the hypothesis of equality in the proportion of subjects violating the two axiom is rejected (McNemar's exact p-value is 0.002 in the PAY treatment and 0.041 in the HYP treatment).

Table 5 confirms that WARP, and therefore the full rationality model, does not describe the data well, especially in the PAY treatment where *less than half* of the subjects fit the model.

Consider now RSM's. The crosstabulation of violations of the two axioms characterizing it is reported in table 6.

Interestingly, in both treatments, no individual who satisfies Expansion violates WARP* (recall that they are logically independent axioms). That is, the (large) number of Expansion violators is not joined by another separate group of WARP* violators in order

[19]For comparisons between the proportion of violations of other pairs of axioms it is not possible to rely on a McNemar test, as violations of either Expansion or WARP* imply violations of WARP (i.e. the relevant contingency table would have structural zeroes). We defer tackling of this issue to our discussion of the relative performance of alternative theories further below.

		PAY				HYP				
		Expansion				Expansion				
		×		✓		×		✓		
		#	%	#	%	#	%	#	%	
WARP*	×	29	*28.4*	0	0	8	*14.3*	0	*0*	× WARP*
	✓	22	*21.6*	51	50	14	*25*	34	*60.7*	✓

Table 6: Violations of WARP* and Expansion

to determine the RSM violators. The RSM violators are simply counted by Expansion violators (of which some are also WARP* violators). The main fact remains, however, that RSM improve only marginally on order maximization in their ability to explain the data for the PAY treatment (decreasing the violations from 52.9 in the case of WARP to 50% in the case of RSM), and they are as bad in the HYP treatment.

Finally, we turn to WARP* and the model of two rationalizability by similarities. From Table 5 we can see that WARP* is satisfied by just below 72% of the subjects in the PAY treatment and just below 86% of the subjects in the HYP treatment.

In summary then:

	PAY		HYP	
	#	%	#	%
Full rationality	48	*47.1*	34	*60.7*
Rational Shortlist Method	51	*50*	34	*60.7*
Two-rationality by similarities	73	*71.6*	48	*85.7*

Table 7: Explanatory power of competing theories.

The three models are nested, that is

$$\text{Full Rationality} \Rightarrow \text{Rational Shortlist Method} \Rightarrow \text{Two-rationality by similarities}$$

In order to compare the incremental 'explanatory' power in each more general theory we take a conservative approach, and look at the 95% exact confidence intervals[20] for the proportion of subjects whose choices are compatible with each theory. In the PAY treatment these confidence intervals are $[0.371, 0.572]$ for the proportion of subjects compatible with Full rationality, $[0.399, 0.601]$ for the proportion of subjects compatible with an RSM and $[0.618, 0.801]$ for the proportion of subjects compatible with two-rationality by similarities. For the HYP treatment the confidence intervals are $[0.467, 0.735]$ for Full rationality and RSM[21] and $[0.738, 0.936]$ for two-rationality by similarities[22]. Thus in both treatments the lower bound of the confidence interval for two-rationality by similarities lies above the upper bound of the other two confidence intervals.

Summary and comment. The general indication we draw from the data is that any model addressing lack of full rationality in a choice function must be able to explain menu effects in the form of Condorcet inconsistency.

This indication is confirmed in the analysis of the three models we have studied in this paper. Neither the full rationality nor the RSM model are compatible with menu effects of the Condorcet inconsistency type, and indeed they both fail badly at explaining the data. The RSM model performs only marginally better than the full rationality model. The proportion of successes in explaining behavior is not increased significantly when weakening WARP to the combination of Expansion and WARP*.

The model of two-rationality by similarities is compatible with Condorcet inconsistency, and it is successful. There is a significant leap in the proportion of successes in explaining behavior when weakening WARP to WARP*. The resulting model can explain 50% more data compared to the other two models, namely over 70% in one treatment and over 85% in the other treatment.

[20]These are computed with the Clopper and Pearson method, which is generally very conservative (the coverage probability can be much greater than the nominal confidence level in small samples). In our case, for the PAY treatment, the sample is large enough and there is no difference between the Clopper-Pearson and the shorter Blyth-Still-Casella exact confidence intervals. For the smaller sample of the HYP treatment, there is a very slight difference between the two methods, as reported below.

[21]The Blyth-Still-Casella confidence interval is $[0.467, 0.728]$.

[22]The Blyth-Still-Casella confidence interval is $[0.747, 0.936]$.

Remember that our test for the 'success' of a model is harsh: we would like each *individual* to satisfy a model in *both* choice contexts (choices among short sequences and choices among long sequences). If instead we focussed separately on the four *choice functions* we have observed, the model of two-rationality by similarities would explain almost 80% of the choice functions in the worst case, and more than 90% in the best case.

3.3 Experimental results II: Other considerations

The data allow a plethora of additional considerations - due to space limitations we cannot analyze them all in this paper, and limit ourselves to highlighting just a few. Two clear patterns concerning violations of the axioms that emerge from both the aggregate data tables and the individual choice data tables are the following:

1. *People are more consistent for longer sequences.* For *each* axiom considered, the proportion of choices or subjects violating it falls as sequence length increases, irrespective of treatment.

2. *People are more consistent if they are not paid.* For *each* axiom considered, the proportion of choices or subjects violating it falls when incentives are removed, i.e. when passing from the PAY to the HYP treatment, regardless of sequence length.

In addition, by looking at the crosstabulation of violations of each axiom by sequence length, we can measure, for each axiom, the *proportion of subjects* failing to satisfy it for at least one choice function. Crosstabulations of this sort allow us to test for each axiom the following:

- *within each treatment*: *(i)* the statistical significance of the fall in the proportions of violations when going from shorter to longer sequences, and *(ii)* whether or not violations observed for different sequence length are associated.

- *across treatments*: whether, controlling for sequence length, the proportion of violations depends on treatment, i.e. whether elicitation of choices by incentive compatible means in the PAY treatment results in a different proportion of subjects violating each axiom as compared to the HYP treatment.

We summarize our main findings in table 8 below (the detailed derivation of this summary is relegated to appendix A.1). Notationwise, π_2 and π_3 refer to the proportions of subjects violating an axiom in choices involving two and three period sequences, respectively, for any given treatment. In addition, for any given sequence length, π_{PAY} and π_{HYP} refer to the proportions of subjects violating an axiom in the PAY and HYP treatment, respectively.

	Within treatment				Across treatment	
	PAY		HYP		$\pi_{PAY} > \pi_{HYP}$	
	$\pi_2 > \pi_3$	random errors	$\pi_2 > \pi_3$	random errors	2 periods	3 periods
CC	✓	×	×	✓	✓	×
PC	×	✓	×	×	✓ (10%)	×
WARP*	✓	× (10%)	×	×	✓	×
EXP	✓ (10%)	×	×	✓	✓	×
WARP	✓	×	×	×	✓	×

Table 8: Comparisons of proportions and association.

In the leftmost part of table 8 (under the heading 'within treatment') we report *(i)* whether or not π_2 is statistically larger than π_3, and *(ii)* whether violation of an axiom for shorter sequences makes it any more likely that the subject violates the same axiom when choosing out of longer sequences, too. If this is not the case, one may assume that differences in the proportions of violations across sequence length are due to the subjects making mistakes independently from one another - in table 8 this lack of association is referred to by the shorthand 'random errors'. In each column, we use a tick (✓) when the relevant statistic is such that the heading in the table 'holds', and a cross (×) when the heading in the column 'fails'. So for point *(i)*, a tick indicates that π_2 is statistically larger than π_3, while a cross indicates that the difference in proportions is not statistically significant.[23] Regarding *(ii)* instead we use a tick to indicate that indeed differences may

[23]To be precise, the null hypothesis of the test for *(i)* is that the proportion of violations is the same regardless of sequence length against a one sided alternative that the proportion of violations for two period sequences is larger than for three period sequences. Then the tick refers to the null being rejected.

be just random, and a cross when this is not the case.[24]

The rightmost part of the table instead reports whether, for any given sequence length, the fall in the proportion of subjects violating a given axiom when moving from the PAY to the HYP treatment is statistically significant (which we denote by a tick ✓) or not (which we denote by a cross $\times$).[25]

There is no clear pattern of association across sequence length for the violations of each axiom ('random error' columns). Broadly, differences in choice behavior between two and three period sequences are more pronounced in the PAY than in the HYP treatment ('$\pi_2 > \pi_3$' columns). Moreover choice behavior for three period sequences does not differ much between the two treatments, whereas it does for choices over two period sequences ('$\pi_{PAY} > \pi_{HYP}$' columns). However, a quick inspection of choice behavior by subject (table 7) shows that for all of the three models analyzed, their ability to explain the data increases in the HYP treatment as compared to the PAY treatment: when moving

---

To test this hypothesis we rely on McNemar' statistic.

[24]To be precise, the null hypothesis of the test is for lack of association between rows and columns in the cross-tabulation (i.e. the odds ratio is equal to 1). If this hypothesis is rejected, then rows and columns are associated, i.e. a subject is much more likely to violate the axiom in choice among three period sequences when he has done so in choice out of two period sequences too. In the table we abuse terminology for the sake of clarity, so that a tick corresponds to a *rejection* of the null hypothesis, while a cross stands for failure to reject. We base this test on Fisher's statistic.

Note that tests *(i)* and *(ii)* are independent, in the sense that a high p-value in the McNemar test does not necessarily imply a high p-value of the Fisher test, and viceversa. For instance, in the table

		long	
		yes	no
short	yes	2	8
	no	1	4

based on the Mc Nemar test one rejects the null of equality of proportion whereas based on the Fisher test one fails to reject the null of lack of association between rows and columns.

[25]To be precise, the null hypothesis of the test is for equality in the proportion of subjects violating a given axiom across the PAY and HYP populations, based on the Fisher test (i.e. the odds ratio for the table with treatments against violation is equal to unity). If this hypothesis is rejected (against the one sided alternative that the proportion of violations in the PAY treatment is larger than in the HYP treatment), then the two proportions are statistically different. In the table we abuse notation for the sake of expositional clarity, so that a tick corresponds to a *rejection* of the null hypothesis, while a cross stands for failure to reject.

from the PAY to the HYP treatment, the percentage of subjects whose choice function is an RSM increases by 10.7%, the percentage of subjects whose choice function can be rationalized in the standard way increases by 13.6%, and the percentage of subjects who are two rational by similarity increases by 14.1%. For the latter two notions of rationalizability these increments are statistically significant,[26] and 'just' not significant for rational shortlist methods.[27]

Comment Our data show a very clear pattern whereby monetary incentives to elicit choices which are the expression of 'true' preferences have the effect of producing less 'rational' behavior. Providing a rigorous explanation for this phenomenon would go beyond the scope of this paper and the bounds of economics. Still, this seems to open a different angle to the discussion on the role of monetary incentives in experiments. In the economics literature this generally revolves around whether or not monetary incentives are necessary to elicit 'true' preferences or the 'best' outcome (see e.g. Camerer and Hogarth [6], Hertwig and Ortmann [14], Read [25] and Harrison and Rutström [7]). However, we note that an empirical regularity in experiments is that subjects are upset when confronted with their own inconsistencies.[28] One might argue that the absence or presence of monetary incentives constitutes a change of (experimental) 'frame', so that what matters is not the composition of the set from which the choice is going to be made, rather the objects it includes *and* whether or not monetary incentives for choice exist. In other words, the choice set when any alternative once chosen is then going to be experienced is a different object from a choice set with the same set of available alternative but where choice itself is just a thought experiment. In addition, there may be other 'external' relevant dimensions to the problem, such as the decision makers' values, motivations and so on, which might influence choice.[29] Based on these considerations, we advance the

[26]In comparing proportions of violations in the PAY and HYP treatments, Fisher test's exact mid-p values are 0.042 for WARP, i.e. standard rationalisability; and 0.022 for WARP*, i.e. two-rationality by similarities.

[27]The exact mid-p value from the Fisher test is is 0.101.

[28]This is based on the casual evidence that generally emerges in de-briefing discussions, although there is psychological literature that deals with the effect of affective states on decisions, see e.g. Luce, Bettman and Payne [19].

[29]This point has been made very clearly by Sen [31].

tentative hypothesis that our results, too, support the position that incentive compatible elicitation of preference is necessary to elicit preferences that are closer to those that a decision maker would display in a real life choice situation. Where choices are only hypothetical in nature, as in the HYP treatment, the decision maker's main concern is that of being consistent, resulting in less frequent violations of the axioms. In this sense, our results invite caution in the use of introspection for testing the 'plausibility' of competing axioms of choice.

3.4 Experimental results III: 'La donna e' mobile'

When looking at violations of the axioms across sexes, a pattern emerges: within each treatment the proportion of women that violate the axioms is higher than men (with one exception). In addition, the pattern we highlighted in the previous section - whereby with the removal of monetary incentives for choice the proportion of violations of our axioms decreases - persists regardless of sex:

	PAY				HYP			
	F		M		F		M	
	#	%	#	%	#	%	#	%
CC	27	*57.4%*	24	*43.6%*	14	*48.3%*	8	*29.6%*
Expansion	27	*57.4%*	24	*43.6%*	14	*48.3%*	8	*29.6%*
WARP*	17	*36.2%*	12	*21.8%*	5	*17.2%*	3	*11.1%*
PC	7	*14.9%*	10	*18.2%*	3	*10.3%*	1	*3.7%*
WARP	27	*57.4%*	27	*49.1%*	14	*48.3%*	8	*29.6%*

Table 9: Violations of the axioms by sex.

The statistical significance of these differences, however, is in general weak. In particular:

1. Within treatment:

    - The difference in proportions of men and women violating Condorcet Consistency and Expansion is statistically significant at 10% confidence level in both the PAY and the HYP treatments;[30]

- The difference in proportions of men and women violating Pairwise Consistency is not significant in either treatment;[31]
- The difference in proportions of men and women violating WARP* is statistically significant at 10% confidence level in the PAY treatment and not statistically significant in the HYP treatment;[32]
- The difference in proportions of men and women violating WARP is not statistically significant in the PAY treatment and statistically significant at 10% confidence level in the HYP treatment.[33]

2. across treatments:

- For female participants, the only difference in the proportions of subjects violating a given axiom across treatments which is statistically significant is for WARP*, for which the Fisher test returns a mid-p value of 0.042;[34]
- For male participants the only differences in proportions which are statistically significant are for Pairwise Consistency, for which the Fisher test yields a p-value of 0.037; and WARP, for which the mid-p value from the Fisher test is 0.051.[35]

The analysis above shows also that there are substantial differences (though not alwasy statistically significant) in the proportion of women and men whose choices conform to either Full rationality (i.e. their choices satisfy WARP) or two-rationality by similarities (i.e. their choices satisfy WARP*). To check differences in the sexes as to the explanatory

[30]The mid-p value for the Fisher test is 0.087 for the PAY treatment and 0.084 for the HYP treatment.

[31]The mid-p value for the Fisher test is 0.336 for the PAY treatment and 0.199 for the HYP treatment.

[32]The mid-p value for the Fisher test is 0.059 for the PAY treatment and 0.27 for the HYP treatment.

[33]The mid-p value for the Fisher test is 0.21 for the PAY treatment and 0.084 for the HYP treatment.

[34]The mid-p value for the Fisher test is equal to 0.224 for Condorcet Consistency, WARP and Expansion, and equal to 0.30 for Pairwise Consistency.

[35]The other mid-p values for the Fisher test are equal to 0.117 for both Condorcet Consistency and Expansion, and to 0.129 for WARP*.

power of RSM we present cross-tabulations of Expansion and WARP* by sex in tables 10 and 11.

PAY

		Females: Expansion × #	Females: Expansion × %	Females: Expansion ✓ #	Females: Expansion ✓ %	Males: Expansion × #	Males: Expansion × %	Males: Expansion ✓ #	Males: Expansion ✓ %		
WARP*	×	17	36.2%	0	0%	12	21.8%	0	0%	×	WARP*
	✓	10	21.3%	20	42.6%	12	21.8%	31	56.4%	✓	

Table 10: RSM by sex in the PAY treatment.

HYP

		Females: Expansion × #	Females: Expansion × %	Females: Expansion ✓ #	Females: Expansion ✓ %	Males: Expansion × #	Males: Expansion × %	Males: Expansion ✓ #	Males: Expansion ✓ %		
WARP*	×	5	17.2%	0	0%	3	11.1%	0	0%	×	WARP*
	✓	9	31.1%	15	51.7%	5	18.5%	19	70.4%	✓	

Table 11: RSM by sex in the HYP treatment.

When considering RSMs, then, the difference across sexes is quite substantial in both treatments (around 13% in the PAY treatment and just short of 20% in the HYP treatment), and it is also statistically significant at 10% confidence level for both treatments.[36] Finally, differences across treatments by sex are not statistically significant.[37]

In summary then:

[36]The mid-p values for the Fisher test of the difference in the proportion of men and women satisfying RSM is equal to 0.087 for the PAY treatment and 0.084 for the HYP treatment.

[37]The mid-p values for the Fisher test of the difference in the proportion of subjects satisfying RSM in the HYP and PAY treatments is equal to 0.228 for Female participants and equal to 0.117 for Male participants.

	PAY				HYP			
	F		M		F		M	
	#	%	#	%	#	%	#	%
Rationalizability	20	*42.6*	28	*50.9*	15	*51.7*	19	*70.4*
Rational Shortlist Methods	20	*42.6*	31	*56.4*	15	*51.7*	19	*70.4*
2-Rationality by Similarities	30	*63.8*	43	*78.2*	24	*82.8*	24	*88.9*

Table 12: Explanatory power of competing theories across sexes

The notion of two-rationality by similarities works better than the other two regardless of sex. Across sexes, there are differences:

- the proportion of men whose choices are rationalizable is higher than the proportion of women in the HYP treatment, but not in the PAY treatment;[38]
- the proportion of men whose choices are RSM is statistically significantly higher than the proportion of women in both the PAY and HYP treatment;[39]
- the proportion of men whose choices are two-rational by similarities is higher than the proportion of women in the PAY but not in the HYP treatment.[40]

It is very hard to explain any of these differences in behavior within a purely economic framework. We leave further analysis to scholars in other fields.

4 Concluding remarks

We have shown in this paper that the standard revealed preference methodology can be successfully used to study 'behavioral' choice procedures. This methodological point has

[38]The Fisher test yields an exact mid-p value equal to 0.215 for the PAY treatment and to 0.084 for the HYP treatment (i.e. for the latter the difference in proportions is significant at 10% confidence level).

[39]The Fisher test yields an exact mid-p value equal to 0.086 for the PAY treatment and to 0.084 for the HYP treatment (i.e. the difference in proportions is significant at 10% confidence level).

[40]The Fisher test yields an exact mid-p value equal to 0.059 for the PAY treatment (i.e. statistical significance is at 10% confidence level) and to 0.272 for the HYP treatment.

also been recently argued by Rubinstein and Salant [29], who inter alia provide a new interpretation of the sequential eliminative heuristics we call Rational Shortlist Methods and provide a different characterization of it.

Both the utility maximization model and Rational Shortlist Methods do not explain well the choice data elicited in our experiment. However, our proposed new model of two-rationality by similarities performs much better. Most violations of utility maximization appear to be due to menu effects (Condorcet inconsistency) rather than to pairwise inconsistency. The main virtue of the model of two-rationality by similarities is its ability to capture in a simple way such menu effects: it is for this reason that it 'outperforms' the other models in the context we have studied.

Though in this paper we have focused on abstract decision making procedures, our experiment is also of specific interest for the theory of choice over time. Although we pursue a more focused analysis of competing theories for the modelling of time preference elsewhere,[41] here we offer a few remarks on this aspect. Choice over time has come under increasing scrutiny in recent years, following a series of observed anomalies that cast doubt on the descriptive validity of the standard model of exponential discounting (see Frederick, Loewenstein and O'Donoghue [10] for a recent and comprehensive survey). Our results suggest that not only the standard model, but also some more recent models that address behavioral anomalies (notably the now very popular hyperbolic discounting model) are descriptively inadequate. No simple change in the functional form of the discounting function will be descriptively adequate, since any such modified theory assumes that choice behavior is based on the maximization of *some* objective function. Neither of the two effects noted in our experiment, pairwise inconsistency and Condorcet inconsistency, can be addressed in this way. In this sense, our results support the arguments put forward by Rubinstein [28].

[41] See Manzini, Mariotti and Mittone [21].

References

[1] Agresti, A. (2002) *Categorical Data Analysis*, Wiley Series in Probability and Statistics, Second Edition, John Wiley and Sons.

[2] Andreoni, J. and W. T. Harbaugh (2006) "Power Indices for Revealed Preferences Tests", mimeo.

[3] Andreoni, J. and J. Miller (2002) "Giving According to GARP: An Experimental Tests of the Consistency of Preferences for Altruism", *Econometrica*, 70: 737-753.

[4] Becker, G. S. (1962) "Irrational Behavior in Economic Theory", *Journal of Political Economy*, 70: 1-13.

[5] Bronars, S. G. (1987) "The Power of Nonparametric Tests of Preference Maximization", *Econometrica*, 55: 693-698.

[6] Camerer, C. F. and R. Hogarth (1999) "The effects of financial incentives in experiments: A review and capital-labor-production framework", *Journal of Risk and Uncertainty*, 19: 7-42.

[7] Harrison, G. and E. Rutström "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods", in C. Plott and V. Smith, eds., *Handbook of Experimental Economics Results*. New York: Elsevier, forthcoming.

[8] Fevrier, P. and M. Visser (2004) "A Study of Consumer Behavior Using Laboratory Data", *Experimental Economics*, 7: 93-114.

[9] Fishburn, P. C. (1970) *Utility theory for decision making*, Wiley, New York.

[10] Frederick, S., G. F. Loewenstein and T. O'Donoghue (2002) "Time discounting and time preferences: a critical review", *Journal of Economic Literature*, 40: 351-401.

[11] Gigerenzer, G., P. Todd, and the ABC Research Group (1999) *Simple Heuristics That Make Us Smart*. New York: Oxford University Press..

[12] Gul, F. and W. Pesendorfer (2001) “Temptation and Self-Control”, *Econometrica*, 69: 1403-1435.

[13] Harbaugh, W. T., K. Krause and T. Berry (2001) “GARP for Kids: On the Development of Rational Choice Behavior”, *American Economic Review*, 91: 1539-1545.

[14] Hertwig, R. and A. Ortmann (2001) ‘Experimental Practices in Economics: A Challenge for Psychologists?’ [target article], *Behavioral and Brain Sciences*, 24: 383-403.

[15] Houy, N. (2006) “Choice Correspondences with States of Mind”, mimeo, Hitotsubashi University.

[16] Kalai, G., A. Rubinstein and R. Spiegler (2002) “Rationalizing Choice Functions by Multiple Rationales”, *Econometrica*, 70, 2481-2488.

[17] Koszegi, B. and M. Rabin “A Model of Reference Dependent Preferences”, forthcoming, *Quarterly Journal of Economics*.

[18] List, J. and C Gallet (2001) “What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?”, *Environmental and Resource Economics*, 20: 241-254.

[19] Luce, M., J. Bettman, and J. Payne (1997) “Choice Processing in Emotionally Difficult Decisions”, *Journal of Experimental Psychology: Learning, Memory and Cognition,* 23:384-405

[20] Manzini, P. and M. Mariotti (2005) “Rationalizing Boundedly Rational Choice: Sequential Rationalizability and Rational Shortlist Methods”, Working paper EconWPA ewp-mic/0407005.

[21] Manzini, P., M. Mariotti and L. Mittone (2006) “Choosing monetary sequences: theory and experimental evidence”, CEEL Working paper 1-06.

[22] Masatlioglu, Y. and E. Ok (2003) “Rational Choice with Status-quo Bias”, *Journal of Economic Theory* 121: 1-29.

[23] Mattei, A. (2000) "Full-Scale Real Tests of Consumer Behavior using Expenditure Data", *Journal of Economic Behavior and Organization*, 43: 487-497.

[24] Moulin, H. (1985) "Choice Functions Over a Finite Set: A Summary" *Social Choice and Welfare*, 2, 147-160.

[25] Read, D. (2005) "Monetary incentives, what are they good for?", *Journal of Economic Methodology*, 12: 265-276.

[26] Roelofsma, P. H. and D. Read, 2000 "Intransitive Intertemporal Choice", *Journal of Behavioral Decision Making*, 13: 161-177.

[27] Rubinstein, A., 1988, "Similarity and Decision Making Under Risk (Is there a utility theory resolution to the Allais Paradox?)", *Journal of Economic Theory*, 46: 145-153.

[28] Rubinstein, A., 2003, "Is it 'Economics and Psychology'? The case of hyperbolic discounting", *International Economic Review*, 44: 1207-1216.

[29] Rubinstein, A. and Y. Salant (2006) "Two Comments on the Principle of Revealed Preference", mimeo, Tel Aviv University, http://arielrubinstein.tau.ac.il/ papers/frames.pdf.

[30] Samuelson, P. (1938) "A Note on the Pure Theory of Consumer's Behavior", *Economica* 5 (1): 61-71.

[31] Sen, A. (1993) "Internal consistency of choice", *Econometrica*, 61: 495-521.

[32] Sen, A. (1997) "Maximization and the Act of Choice", *Econometrica*, 65, 745-779.

[33] Sippel, R. (1997) "An Experiment on the Pure Theory of Consumer's Behavior", *Economic Journal*, 107: 1431-1444.

[34] Suzumura, K. (1983) *Rational Choice, Collective Decisions, and Social Welfare*, Cambridge University Press, Cambridge U.K.

[35] Tversky, A. (1969) "Intransitivity of Preferences", *Psychological Review*, 76: 31-48.

[36] Tversky, A. (1972) “Elimination By Aspects: A Theory of Choice”, *Psychological Review*, 79: 281-299.

[37] Varian, H. (2005) “Revealed Preference”, mimeo, University of California at Berkeley, http://www.sims.berkeley.edu/~hal/Papers/2005/revpref.pdf.

[38] Waite, T. (2001) “Intransitive preferences in hoarding gray jays (Perisoreus canadensis)”, *Behavioral Ecology and Sociobiology*, 50: 116 - 121.

[39] Yee, M., E. Dahan, J. R. Hauser and J. Orlin (2005) “Greedoid-Based Non-Compensatory Inference”, *Journal of Marketing Science*, forthcoming.

A Appendices

A.1 Analysis of violations of the axioms, crosstabulated by sequence length

In this section we report in full the crosstabulation of violations of each of the five axioms (Condorcet Consistency, Pairwise Consistency, WARP, Expansion and WARP*) considered in the main text by sequence length, on which the summary tables in the main text are based. This allows us to measure for each axiom the proportion of subjects failing to satisfy it for either longer or shorter sequences.

In the tables that follow we use a cross ($\times$) to indicate that the axiom is violated and a tick ($\checkmark$) to indicate that it holds. For each axiom we report: *(i)* within each treatment (i.e. PAY or HYP) whether the proportion of subjects violating the axiom falls with the increase in sequence length in a statistically significant measure, and whether violations for shorter sequences are associated to violations for longer sequences (i.e. ‘random mistakes’ in the sense of section 3.3); and *(ii)* across treatment whether monetary incentives have an effect on the proportions of subjects violating each axiom.

Condorcet Consistency

		PAY				HYP			
		3 periods				3 periods			
		×		✓		×		✓	
		#	%	#	%	#	%	#	%
2 periods	×	18	*17.6*	21	*20.6*	4	*7.1*	10	*17.9*
	✓	12	*11.8*	**51**	*50*	8	*14.3*	**34**	*60.7*

Table 13: Violations of Condorcet Consistency for different sequence length.

- *Within treatment: PAY.* For the PAY treatment, the proportion of subjects violating Condorcet Consistency falls from 38.2% (i.e. 17.6 + 20.6) to 29.4% (i.e. 17.6 + 11.8) as sequence length increases, and this difference is statistically significant at 10% confidence level (the exact p-value for the McNemar test is 0.081). In addition, we reject the hypothesis of lack of association between violations of Condorcet Consistency in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.002) - in short, we can reject the hypothesis that violations are due to random mistakes.

- *Within treatment: HYP.* For the HYP treatment, the proportion of subjects violating Condorcet Consistency falls slightly from 25% to 21.4% as sequence length increases, but this difference is not statistically significant (the exact p-value of the McNemar test is 0.407). In addition, we cannot reject the hypothesis of lack of association between violations of Condorcet Consistency in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.237) - in short, we cannot reject the hypothesis that violations are due to random mistakes.

- *Across treatment: two period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations of Condorcet Consistency falls when moving from the PAY (38.2%) to the HYP (25%) treatment, and this difference is statistically significant (Fisher test yields an exact mid-p value of 0.047);

- *Across treatment: three period sequences.* Comparing across treatments we note that for the three period sequences the percentage of violations of Condorcet consistency falls when moving from the PAY (29.4%%) to the HYP (21.4%) treatment, and this difference is not statistically significant (Fisher test yields an exact mid-p value of 0.14).

Pairwise Consistency

		PAY				HYP			
		3 periods				3 periods			
		×		✓		×		✓	
		#	%	#	%	#	%	#	%
2 periods	×	2	*1.9*	10	*9.9*	2	*3.6*	1	*1.8*
	✓	5	*4.9*	**85**	*83.3*	1	*1.8*	**52**	*92.8*

Table 14: Violations of Pairwise Consistency for different sequence length.

- *Within treatment: PAY.* For the PAY treatment, the proportion of subjects violating Pairwise Consistency falls from 11.8% to 6.8% as sequence length increases, but this difference is not statistically significant (the exact p-value for the McNemar test is 0.151). In addition, we cannot reject the hypothesis of lack of association between violations of Pairwise Consistency in choicess over two and three period sequences (the exact mid-p value for the Fisher test is 0.113) - in short, we cannot reject the hypothesis that violations are due to random mistakes.

- *Within treatment: HYP.* For the HYP treatment, the proportion of subjects violating Pairwise Consistency stays unchanged at 5.4%. In addition, we reject the hypothesis of lack of association between violations of Pairwise Consistency in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.003) - in short, we reject the hypothesis that violations are due to random mistakes.

- *Across treatment: two period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations of Pairwise Consistency falls when moving from the PAY (11.8%) to the HYP (5.4%) treatment, and this difference is statistically significant at 10% confidence level (Fisher test yields an exact mid-p value of 0.100);

- *Across treatment: two period sequences.* Comparing across treatments we note that for the three period sequences the percentage of violations of Pairwise Consistency falls when moving from the PAY (6.9%) to the HYP (5.4%) treatment, and this difference is not statistically significant (Fisher test yields an exact mid-p value of 0.375).

WARP

		PAY				HYP			
		3 periods				3 periods			
		×		✓		×		✓	
		#	%	#	%	#	%	#	%
2 periods	×	19	*18.6*	24	*23.5*	7	*12.5*	9	*16.1*
	✓	11	*10.8*	**48**	*47.1*	6	*10.7*	**34**	*60.7*

Table 15: Violations of WARP for different sequence length.

- *Within treatment: PAY.* For the PAY treatment, the proportion of subjects violating WARP falls from 42.1% to 29.4% as sequence length increases, and this difference is statistically significant (the exact p-value for the McNemar test is 0.020). In addition, we reject the hypothesis of lack of association between violations of WARP in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.003) - in short, we can reject the hypothesis that violations are due to random mistakes.

- *Within treatment: HYP.* For the HYP treatment, the proportion of subjects violating WARP falls from 28.6% to 23.2% as sequence length increases, but this difference is not statistically significant (the exact p-value of the McNemar test is 0.304). In addition, we reject the hypothesis of lack of association between violations of WARP in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.017) - in short, we reject the hypothesis that violations are due to random mistakes.

- *Across treatment: two period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations of WARP falls when moving from the PAY (42.1%) to the HYP (28.6%) treatment, and this difference is statistically significant (Fisher test yields an exact mid-p value of 0.047);
- *Across treatment: three period sequences.* Comparing across treatments we note that for the three period sequences the percentage of violations of WARP falls when moving from the PAY (29.4%) to the HYP (23.2%) treatment, but this difference is not statistically significant (Fisher test yields an exact mid-p value of 0.206).

Expansion

		PAY				HYP			
		3 periods				3 periods			
		×		✓		×		✓	
		#	%	#	%	#	%	#	%
2 periods	×	18	*17.6*	21	*20.6*	5	*8.9*	9	*16.1*
	✓	12	*11.8*	**51**	*50*	8	*14.3*	**34**	*60.7*

Table 16: Violations of Expansion for different sequence length.

- *Within treatment: PAY.* For the PAY treatment, the proportion of subjects violating Expansion falls from 38.2% to 29.4% as sequence length increases, and this difference is statistically significant at 10% confidence level (the exact p-value for the McNemar test is 0.081). In addition, we reject the hypothesis of lack of association between violations of Expansion in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.002) - in short, we can reject the hypothesis that violations are due to random mistakes.

- *Within treatment: HYP.* For the HYP treatment, the proportion of subjects violating Expansion falls from 25% to 23.2% as sequence length increases, but this difference is not statistically significant (the exact p-value of the McNemar test is 0.5). In addition, we cannot reject the hypothesis of lack of association between violations of Expansion in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.117) - in short, we cannot reject the hypothesis that violations are due to random mistakes.

- *Across treatment: two period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations of Expansion falls when moving from the PAY (38.2%) to the HYP (25%) treatment, and this difference is statistically significant (Fisher test yields an exact mid-p value of 0.047);

- *Across treatment: three period sequences.* Comparing across treatments we note that for the three period sequences the percentage of violations of Expansion falls when moving from the PAY (29.4%) to the HYP (23.2%) treatment, but this difference is not statistically significant (Fisher test yields an exact mid-p value of 0.206).

WARP*

		PAY				HYP			
		3 periods				3 periods			
		×		✓		×		✓	
		#	%	#	%	#	%	#	%
2 periods	×	5	*4.9*	17	*16.7*	3	*5.4*	3	*5.4*
	✓	7	*6.8*	**73**	*71.6*	2	*3.6*	**48**	*85.6*

Table 17: Violations of WARP* for different sequence length.

- *Within treatment: PAY.* For the PAY treatment, the proportion of subjects violating WARP* falls from 21.6% to 11.7% as sequence length increases, and this difference is statistically significant (the exact p-value for the McNemar test is 0.032). In addition, we can reject at 10% confidence level the hypothesis of lack of association between violations of WARP* in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.051) - in short, we can reject the hypothesis that violations are due to random mistakes.

- *Within treatment: HYP.* For the HYP treatment, the proportion of subjects violating WARP* falls from 10.8% to 9% as sequence length increases, but this difference is not statistically significant (the exact p-value of the McNemar test is 0.5). In addition, we reject the hypothesis of lack of association between violations of WARP* in choices over two and three period sequences (the exact mid-p value for the Fisher test is 0.003) - in short, we reject the hypothesis that violations are due to random mistakes.

- *Across treatment: two period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations fof WARP* alls when moving from the PAY (21.6%) to the HYP (10.8%) treatment, and this difference is statistically significant (Fisher test yields an exact mid-p value of 0.045);

- *Across treatment: three period sequences.* Comparing across treatments we note that for the two period sequences the percentage of violations of WARP* falls slightly when moving from the PAY (11.7%) to the HYP (9%) treatment, and this difference is statistically significant (Fisher test yields an exact mid-p value of 0.298).

	PAY		HYP	
	#	%	#	%
Condorcet Consistency	51	*50*	22	*39.3*
Expansion	51	*50*	22	*39.3*
WARP*	29	*28.4*	8	*14.4*
Pairwise Consistency	17	*16.7*	4	*7.2*
WARP	54	*52.9*	22	*39.3*

Table 18: Overall axiom violations.

Based on tables 16-17 we can summarize the overall violations of the axioms considered *by experimental subject* in table 18, whereas the outcome of the various tests are summarized in table 8 in the main text.

A.2 Failures of Rational Shortlist Methods by sequence length

We report in tables 19 and 20 the crosstabulation of violations of Expansion and WARP* for each choice function in each treatment.

Rational Shortlist Methods: PAY

Sequences	WARP*	Expansion × #	Expansion × %	Expansion ✓ #	Expansion ✓ %
Two period sequences	×	22	*21.6*	0	*0*
Two period sequences	✓	17	*16.7*	**63**	*61.7*
Three period sequences	×	12	*11.8*	0	*0*
Three period sequences	✓	18	*17.6*	**72**	*70.6*

Table 19: Violations of WARP* and Expansion by sequence length, PAY treatment

Rational Shortlist Methods: HYP

Sequences	WARP*	Expansion × #	Expansion × %	Expansion ✓ #	Expansion ✓ %
Two period sequences	×	6	*10.7*	0	0
Two period sequences	✓	8	*14.3*	42	75
Three period sequences	×	5	14.3	0	*0*
Three period sequences	✓	8	8.9	**43**	*76.8*

Table 20: Violations of WARP* and Expansion by sequence length, HYP treatment

In addition, we also distinguish more finely the number of subjects which violate which axioms in which choice function in tables 21 and 22. In this way we can see that, although there is no theoretical reason for this to happen, the fact that violations of WARP* imply violations of Expansion is an empirical regularity.

Rational Shortlist Methods: PAY

2\3→ ↓	**both violated**		**EXP only**		**WARP* only**		**none**	
	#	%	#	%	#	%	#	%
both violated	5	*4.9*	8	*7.8*	0	*0*	9	*8.8*
EXP only	3	*2.9*	2	*1.9*	0	*0*	12	*11.8*
WARP* only	0	*0*	0	*0*	0	*0*	0	*0*
none violated	4	*3.9*	8	*7.8*	0	*0*	51	*50*

Table 21: Violations of the axioms characterising RSM by sequence length - PAY.

Rational Shortlist Methods: HYP

2\3→ ↓	**both violated**		**EXP only**		**WARP* only**		**none**	
	#	%	#	%	#	%	#	%
both violated	3	5.4	0	*0*	0	*0*	3	*5.4*
EXP only	0	0	2	*3.6*	0	*0*	6	*10.7*
WARP* only	0	0	0	*0*	0	*0*	0	*0*
none violated	2	3.6	6	*10.7*	0	*0*	34	*60.6*

Table 22: Violations of the axioms characterising RSM by sequence length - HYP.

A.3 Instructions

Please note: you are not allowed to communicate with the other participants for the entire duration of the experiment.

The instructions are the same for all you. You are taking part in an experiment to study intertemporal preferences. The project is financed by the ESRC.

Shortly you will see on your screen a series of displays. Each display contains various remuneration plans worth the same total amount of 48 Euros each, staggered in three, six and nine months installments. For every display you will have to select the plan that you prefer, clicking on the button with the letter corresponding to the chosen plan. (HYP: These remuneration plans are purely hypothetical. At the end of the experiment you'll be given a participation fee of 5 Euros.) (PAY: At the end of the experiment one of the displays will be drawn at random and your remuneration will be made according to the plan you have chosen in that display).

In order to familiarize yourself with the way the plans will be presented on the screen, we shall now give you a completely hypothetical example, based on a total remuneration of 7 Euros.

Plan A

How much	When
3 Euros	in one year
1 Euro	in two years
1 Euro	in three years
2 Euro	in four years

Plan B

How much	When
1 Euro	in one year
2 Euros	in two years
3 Euros	in three years
1 Euro	in four years

In this example plan A yields 7 Euros in total in installments of 3 Euros, 1 Euro, 1 Euro and 2 Euros in a year, two years, three years and four years from now, respectively, while plan B yields 7 Euros in total in installments of 1 Euro, 2 Euros, 3 Euros and 1 Euro in a year, two years, three years and four years from now, respectively.

A.4 Raw Data

We describe below the variable names used in Table 23:

T: treatment (0 for PAY and 1 for HYP)

SS: session number

SB: subject number

SX: subject's sex (F for Female and M for Male)

Choices between plans are coded as follows: *abcdn* indicates the choice between plans a, b, c and d of length n periods. A value of 1, 2, 3 or 4 indicates that a, b, c or d, respectively, was chosen. Similar for choices *abcn* (involving three plans only) and *abn* (involving two plans only).

T	SS	SB	SX	ki3	id3	dk3	ij3	jk3	jd3	kid3	jki3	djk3	idj3	kidj3	ki2	id2	dk2	ij2	jk2	jd2	kid2	jki2	djk2	idj2	kidj2
0	1	0	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	1	1	F	1	1	2	1	1	1	1	2	3	2	1	1	2	2	2	2	2	1	2	1	3	4
0	1	2	F	2	1	2	1	2	2	2	2	2	1	1	1	1	1	1	2	2	1	3	3	1	1
0	1	3	F	1	1	2	1	2	1	1	3	3	1	1	1	1	2	1	2	2	1	2	3	1	1
0	1	4	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	1	5	F	1	2	1	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	1	6	M	1	2	1	1	2	2	3	3	1	2	3	1	2	1	1	1	2	3	2	1	2	3
0	1	7	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	1	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	1	9	F	2	1	2	2	1	1	2	1	2	3	4	2	1	2	1	1	1	2	3	2	3	4
0	2	0	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	1	F	2	2	1	1	1	2	3	3	1	2	4	1	1	2	1	2	2	1	2	1	2	1
0	2	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	3	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	4	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	2	5	F	1	1	2	1	2	2	3	2	3	1	1	1	1	2	1	2	2	1	2	3	2	1
0	2	6	M	1	2	1	1	2	2	1	2	1	2	1	1	2	1	1	2	2	3	2	1	2	3
0	2	7	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3

0	2	9	M	2	1	2	1	2	1	2	2	3	1	2	1	1	2	1	2	1	1	3	3	1	2
0	2	10	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	11	F	1	2	1	1	2	2	3	2	1	2	3	1	1	1	1	2	2	1	2	1	3	3
0	2	12	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	13	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	14	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	2	15	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	3	0	M	2	2	2	1	2	2	2	3	1	1	1	2	1	2	1	2	2	2	3	2	1	1
0	3	1	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	3	2	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	3	2	3
0	3	3	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	3	4	M	1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	2	2	1	2	3	2	1
0	3	5	M	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	3	2	1	2	3
0	3	6	F	1	2	2	1	1	2	3	2	3	2	1	1	2	1	1	2	2	1	2	1	2	3
0	3	7	F	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	1	2	1	2	3
0	3	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	0	F	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	1	2	1	2	1
0	4	1	M	1	2	1	2	2	2	3	2	1	2	4	1	2	1	1	2	2	3	2	1	2	3
0	4	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	3	M	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	1	2	3	2	1
0	4	4	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	1	2	3	2	1	2	3
0	4	5	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	6	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	7	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	9	M	1	1	2	1	2	1	1	2	3	3	1	1	1	2	1	1	1	1	2	3	1	1
0	4	10	M	1	2	1	1	2	2	3	3	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	4	11	F	1	1	2	1	2	2	1	2	3	1	1	1	1	1	1	2	2	1	3	3	1	3
0	5	0	F	1	1	1	2	2	2	1	3	1	3	2	1	1	1	2	2	2	3	3	1	3	1
0	5	1	F	1	1	2	1	2	1	2	3	2	1	2	2	1	2	2	1	1	1	3	2	1	4

0	5	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	5	3	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	5	4	F	1	2	1	1	2	2	3	2	1	2	1	1	2	1	1	2	2	2	2	3	2	1
0	5	5	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	5	6	M	1	2	1	1	2	2	1	2	1	2	1	1	2	1	1	2	2	1	3	1	2	3
0	5	7	F	1	1	2	1	2	1	1	2	3	1	1	2	1	2	1	2	1	2	3	3	1	2
0	6	0	F	1	2	1	1	2	2	1	2	2	2	3	1	2	1	2	2	2	3	2	1	2	3
0	6	1	M	1	2	1	2	1	1	3	1	2	2	4	2	2	1	2	1	2	3	1	1	2	3
0	6	2	F	1	2	1	1	2	2	3	2	1	2	3	1	2	2	1	2	2	3	2	1	2	3
0	6	3	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	6	4	M	1	1	2	1	2	1	1	2	3	1	1	1	1	2	2	2	2	1	2	3	1	1
0	6	5	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	6	6	M	1	1	1	1	1	1	3	2	3	2	1	1	1	2	1	1	2	3	2	1	3	1
0	6	7	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	6	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	6	9	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	6	10	F	1	2	1	1	2	2	3	2	1	2	3	1	2	2	1	2	2	3	2	1	2	3
0	6	11	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	3	2	3
0	7	0	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	7	1	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	7	2	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	7	3	F	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	3	2	1	2	1
0	7	4	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	1	2	3
0	7	5	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	7	6	F	1	2	2	1	2	2	1	2	3	2	1	1	2	1	1	2	2	3	2	3	2	1
0	7	7	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
0	7	8	F	2	1	2	1	2	1	2	3	3	1	2	2	1	2	1	2	2	2	3	3	1	2
0	7	9	M	1	2	2	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	7	10	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	7	11	M	1	2	1	2	1	2	3	3	1	2	3	1	2	1	1	2	2	3	2	1	2	3

0	8	0	M	1	1	2	1	1	1	2	3	1	1	4	2	1	2	1	1	1	2	1	2	3	4
0	8	1	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	8	2	M	1	2	1	2	2	2	3	2	1	2	3	1	2	1	2	1	1	1	1	2	3	4
0	8	3	F	1	2	1	1	2	2	1	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	8	4	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	8	5	M	1	2	2	1	2	1	1	1	1	1	3	1	2	2	2	2	1	3	2	1	2	4
0	8	6	F	1	1	2	1	2	2	1	2	3	1	1	1	1	2	1	2	2	1	2	3	2	1
0	8	7	M	1	2	1	2	2	2	1	2	1	2	3	1	2	1	1	2	2	3	2	2	2	3
0	8	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	8	9	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	3	3	1	2	3
0	8	10	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	9	0	M	1	2	2	2	2	2	1	2	3	3	1	2	2	2	1	2	2	3	2	3	2	1
0	9	1	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	9	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	9	3	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	9	4	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
0	9	5	M	1	2	1	1	2	2	3	2	1	2	3	1	2	2	1	1	2	1	2	3	2	1
0	9	6	F	2	1	2	1	1	1	2	3	2	1	2	1	1	2	1	1	2	2	1	3	1	4
0	9	7	M	1	2	1	1	2	2	3	2	1	2	1	1	2	2	1	2	2	3	2	3	2	1
0	9	8	M	1	1	2	1	2	1	1	2	3	1	1	1	1	2	1	2	1	2	2	3	1	1
0	9	9	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	3	2	3
0	9	10	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	1	2	1
0	9	11	M	1	1	2	2	2	1	1	2	3	3	1	1	1	2	1	2	1	1	2	3	3	1
1	1	0	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	1	F	1	1	2	1	2	2	1	2	3	1	1	1	2	2	1	2	2	1	2	3	1	1
1	1	2	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	3	F	1	1	2	1	2	2	1	2	3	1	1	1	1	2	1	2	2	1	2	3	1	1
1	1	4	F	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
1	1	5	F	1	1	2	1	2	2	1	2	3	1	1	1	2	2	1	2	2	1	2	3	2	1
1	1	6	M	2	1	2	1	2	2	2	3	3	1	2	2	1	2	1	2	2	2	3	3	1	2

1	1	7	F	1	2	1	1	2	2	1	2	3	2	3	1	2	2	1	2	2	1	2	3	2	1
1	1	8	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	9	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	10	F	1	1	1	1	2	2	3	2	3	1	1	1	2	1	1	1	2	1	2	1	1	1
1	1	12	F	1	2	1	1	2	2	3	3	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	13	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	14	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	1	15	F	1	2	2	1	2	2	3	2	3	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	0	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	1	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	3	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	4	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	5	M	1	1	1	2	2	2	1	1	2	3	1	2	2	2	1	2	1	3	3	3	2	3
1	2	6	M	1	1	2	1	2	1	1	2	3	1	1	1	1	2	1	2	2	1	2	3	1	1
1	2	7	F	1	2	1	1	2	2	1	2	3	2	1	1	2	1	1	2	2	1	2	3	2	1
1	2	8	M	2	1	2	1	1	1	2	3	2	1	2	2	1	2	1	2	1	2	3	3	1	2
1	2	9	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	2	10	F	1	2	2	1	2	2	1	2	3	2	3	1	2	1	1	2	2	3	2	1	2	3
1	3	0	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	3	1	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	3	2	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	1	2	1	2	3
1	3	3	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	3	4	M	2	2	1	1	1	2	3	3	1	2	3	1	1	2	1	2	2	1	2	3	1	1
1	3	5	F	1	1	2	1	2	2	1	2	3	2	3	1	2	2	1	1	2	1	2	3	2	1
1	3	6	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	3	7	F	2	1	2	2	1	1	2	1	2	3	4	2	1	2	2	1	1	2	1	2	3	4
1	3	8	M	1	1	2	1	2	2	1	3	3	2	1	1	2	1	1	2	2	2	3	3	2	3
1	3	9	M	2	1	2	1	1	1	2	3	2	1	2	2	1	2	1	1	1	2	3	2	1	2
1	3	10	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3

1	3	11	F	1	1	2	1	2	2	1	2	3	1	1	1	2	1	1	2	2	1	2	1	2	1
1	3	12	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	3	1	2	2
1	3	13	M	1	2	2	1	2	2	1	3	3	1	3	2	2	1	1	1	2	2	1	2	1	2
1	3	14	F	1	1	2	1	2	1	1	2	3	1	1	1	1	2	1	2	2	1	2	3	1	1
1	3	15	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	4	0	M	2	1	2	2	1	1	2	1	2	3	4	2	1	2	1	1	1	2	3	3	1	2
1	4	1	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	4	2	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	4	3	M	1	1	2	2	2	1	1	2	3	1	1	1	1	2	1	2	2	1	2	3	1	1
1	4	4	F	2	2	2	1	2	1	1	3	3	2	2	2	1	2	1	2	1	2	3	1	3	2
1	4	5	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	4	6	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	3	1
1	4	7	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	3	2	3
1	4	8	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	1	2	1	2	3
1	4	9	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	3	2	1
1	4	10	M	1	2	2	1	2	2	1	2	3	2	1	1	2	2	1	2	2	1	2	1	2	3
1	4	11	F	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3
1	4	12	F	1	2	1	1	2	2	3	2	1	2	1	1	2	1	1	2	2	3	2	1	2	3
1	4	13	M	1	2	1	1	2	2	3	2	1	2	3	1	2	1	1	2	2	3	2	1	2	3

Table 23: Raw data.

A.5 Frequency distribution of choice profiles

Below we present the frequency distribution for the choice functions we have observed with both two and three period sequences (Tables 24 and 25). How should one read these tables? Because it would be impractical to list all the $20,736$ possible choice functions, in Tables 24 and 25 each choice function is coded in the format X-Y, where X refers to binary choice profiles and Y refers to non binary choice profiles. There are 64 possible combinations of choices from binary sets, so X is a number between 1 and 64, with the corresponding choice profiles listed in Table 26. As for choices out of the non-binary sets, there are 324 possible combinations, so Y is a number between 1 and 324, with the corresponding profiles listed in Table 27.

For instance, consider the modal choice profile in both tables 24 and 25, 51-195. From Table 26 one can see that profile 51 corresponds to $KI = 0$, $ID = 1$, $DK = 0$, $IJ = 0$, $JK = 1$ and $JD = 1$, while from Table 27 profile 195 corresponds to $KID = 3$, $JKI = 2$, $DJK = 1$, $IDJ = 2$, $KIDJ = 3$. Thus the corresponding choice function is $\gamma(\{K, I\}) = K$, $\gamma(\{I, D\}) = D$, $\gamma(\{D, K\}) = D$, $\gamma(\{I, J\}) = I$, $\gamma(\{J, K\}) = K$, $\gamma(\{J, D\}) = D$, $\gamma(\{K, I, D\}) = D$, $\gamma(\{J, K, I\}) = K$, $\gamma(\{D, J, K\}) = D$ and $\gamma(\{K, I, D, J\}) = D$.

two period sequences	PAY	HYP
5-22	1	0
6-98	0	1
6-107	0	1
6-308	1	0
6-314	1	0
11-307	1	0
14-259	1	0
14-308	0	1
21-23	1	0
21-76	1	0
21-106	1	0
22-107	1	1
22-143	0	1

24-216	0	1
31-276	1	0
35-4	0	1
35-195	2	0
36-92	0	1
37-60	1	0
37-263	1	0
39-49	1	1
44-192	1	0
49-25	1	0
49-187	1	0
49-220	1	0
51-31	1	1
51-33	1	0
51-49	1	1
51-50	1	0
51-51	1	0
51-117	0	1
51-193	2	2
51-195	47	27
51-196	1	0
51-204	1	0
51-213	3	1
51-215	0	1
53-22	1	5
53-31	1	0
53-49	2	0
54-17	1	0
54-107	1	1

55-22	0	1
55-31	1	0
55-49	7	4
55-51	1	0
55-76	0	1
55-193	1	1
55-195	2	0
55-198	1	0
56-51	1	0
57-63	1	0
59-195	1	0
61-22	1	0
63-301	1	0
Total	102	56

Table 24: Frequency distribution of choice functions for two period sequences.

tree period sequences	PAY	HYP
1-51	1	0
5-49	1	0
5-251	1	0
6-98	1	2
11-282	1	0
14-308	1	2
21-22	3	2
21-25	1	0
21-76	1	0
21-98	1	0
22-104	1	0

22-107	1	0
23-163	1	0
24-133	0	1
29-22	0	1
29-76	1	0
36-198	0	1
36-279	1	0
39-51	1	0
43-198	1	0
49-24	0	1
51-31	3	0
51-33	2	1
51-49	1	1
51-193	1	0
51-195	48	27
51-198	2	1
51-202	1	0
51-211	0	1
53-22	2	4
53-24	1	0
53-52	0	1
53-211	0	1
54-14	1	0
54-107	0	1
55-49	14	4
55-187	0	1
55-195	1	0
55-211	0	1
55-213	0	1

56-8	1	0
57-64	0	1
57-142	1	0
59-193	1	0
59-195	1	0
59-276	1	0
63-76	1	0
Total	102	56

Table 25: Frequency distribution of choice functions for three period sequences.

profile of binary choices	KI	ID	DK	IJ	JK	JD
1	0	0	0	0	0	0
2	1	0	0	0	0	0
3	0	1	0	0	0	0
4	1	1	0	0	0	0
5	0	0	1	0	0	0
6	1	0	1	0	0	0
7	0	1	1	0	0	0
8	1	1	1	0	0	0
9	0	0	0	1	0	0
10	1	0	0	1	0	0
11	0	1	0	1	0	0
12	1	1	0	1	0	0
13	0	0	1	1	0	0
14	1	0	1	1	0	0
15	0	1	1	1	0	0
16	1	1	1	1	0	0
17	0	0	0	0	1	0

18	1	0	0	0	1	0
19	0	1	0	0	1	0
20	1	1	0	0	1	0
21	0	0	1	0	1	0
22	1	0	1	0	1	0
23	0	1	1	0	1	0
24	1	1	1	0	1	0
25	0	0	0	1	1	0
26	1	0	0	1	1	0
27	0	1	0	1	1	0
28	1	1	0	1	1	0
29	0	0	1	1	1	0
30	1	0	1	1	1	0
31	0	1	1	1	1	0
32	1	1	1	1	1	0
33	0	0	0	0	0	1
34	1	0	0	0	0	1
35	0	1	0	0	0	1
36	1	1	0	0	0	1
37	0	0	1	0	0	1
38	1	0	1	0	0	1
39	0	1	1	0	0	1
40	1	1	1	0	0	1
41	0	0	0	1	0	1
42	1	0	0	1	0	1
43	0	1	0	1	0	1
44	1	1	0	1	0	1
45	0	0	1	1	0	1
46	1	0	1	1	0	1

47	0	1	1	1	0	1
48	1	1	1	1	0	1
49	0	0	0	0	1	1
50	1	0	0	0	1	1
51	0	1	0	0	1	1
52	1	1	0	0	1	1
53	0	0	1	0	1	1
54	1	0	1	0	1	1
55	0	1	1	0	1	1
56	1	1	1	0	1	1
57	0	0	0	1	1	1
58	1	0	0	1	1	1
59	0	1	0	1	1	1
60	1	1	0	1	1	1
61	0	0	1	1	1	1
62	1	0	1	1	1	1
63	0	1	1	1	1	1
64	1	1	1	1	1	1

Table 26: Possible profiles in binary choice.

profile	KID	JKI	DJK	IDJ	KIDJ
1	1	1	1	1	1
2	2	1	1	1	1
3	3	1	1	1	1
4	1	2	1	1	1
5	2	2	1	1	1
6	3	2	1	1	1
7	1	3	1	1	1
8	2	3	1	1	1

9	3	3	1	1	1
10	1	1	2	1	1
11	2	1	2	1	1
12	3	1	2	1	1
13	1	2	2	1	1
14	2	2	2	1	1
15	3	2	2	1	1
16	1	3	2	1	1
17	2	3	2	1	1
18	3	3	2	1	1
19	1	1	3	1	1
20	2	1	3	1	1
21	3	1	3	1	1
22	1	2	3	1	1
23	2	2	3	1	1
24	3	2	3	1	1
25	1	3	3	1	1
26	2	3	3	1	1
27	3	3	3	1	1
28	1	1	1	2	1
29	2	1	1	2	1
30	3	1	1	2	1
31	1	2	1	2	1
32	2	2	1	2	1
33	3	2	1	2	1
34	1	3	1	2	1
35	2	3	1	2	1
36	3	3	1	2	1
37	1	1	2	2	1

38	2	1	2	2	1
39	3	1	2	2	1
40	1	2	2	2	1
41	2	2	2	2	1
42	3	2	2	2	1
43	1	3	2	2	1
44	2	3	2	2	1
45	3	3	2	2	1
46	1	1	3	2	1
47	2	1	3	2	1
48	3	1	3	2	1
49	1	2	3	2	1
50	2	2	3	2	1
51	3	2	3	2	1
52	1	3	3	2	1
53	2	3	3	2	1
54	3	3	3	2	1
55	1	1	1	3	1
56	2	1	1	3	1
57	3	1	1	3	1
58	1	2	1	3	1
59	2	2	1	3	1
60	3	2	1	3	1
61	1	3	1	3	1
62	2	3	1	3	1
63	3	3	1	3	1
64	1	1	2	3	1
65	2	1	2	3	1
66	3	1	2	3	1

67	1	2	2	3	1
68	2	2	2	3	1
69	3	2	2	3	1
70	1	3	2	3	1
71	2	3	2	3	1
72	3	3	2	3	1
73	1	1	3	3	1
74	2	1	3	3	1
75	3	1	3	3	1
76	1	2	3	3	1
77	2	2	3	3	1
78	3	2	3	3	1
79	1	3	3	3	1
80	2	3	3	3	1
81	3	3	3	3	1
82	1	1	1	1	2
83	2	1	1	1	2
84	3	1	1	1	2
85	1	2	1	1	2
86	2	2	1	1	2
87	3	2	1	1	2
88	1	3	1	1	2
89	2	3	1	1	2
90	3	3	1	1	2
91	1	1	2	1	2
92	2	1	2	1	2
93	3	1	2	1	2
94	1	2	2	1	2
95	2	2	2	1	2

96	3	2	2	1	2
97	1	3	2	1	2
98	2	3	2	1	2
99	3	3	2	1	2
100	1	1	3	1	2
101	2	1	3	1	2
102	3	1	3	1	2
103	1	2	3	1	2
104	2	2	3	1	2
105	3	2	3	1	2
106	1	3	3	1	2
107	2	3	3	1	2
108	3	3	3	1	2
109	1	1	1	2	2
110	2	1	1	2	2
111	3	1	1	2	2
112	1	2	1	2	2
113	2	2	1	2	2
114	3	2	1	2	2
115	1	3	1	2	2
116	2	3	1	2	2
117	3	3	1	2	2
118	1	1	2	2	2
119	2	1	2	2	2
120	3	1	2	2	2
121	1	2	2	2	2
122	2	2	2	2	2
123	3	2	2	2	2
124	1	3	2	2	2

125	2	3	2	2	2
126	3	3	2	2	2
127	1	1	3	2	2
128	2	1	3	2	2
129	3	1	3	2	2
130	1	2	3	2	2
131	2	2	3	2	2
132	3	2	3	2	2
133	1	3	3	2	2
134	2	3	3	2	2
135	3	3	3	2	2
136	1	1	1	3	2
137	2	1	1	3	2
138	3	1	1	3	2
139	1	2	1	3	2
140	2	2	1	3	2
141	3	2	1	3	2
142	1	3	1	3	2
143	2	3	1	3	2
144	3	3	1	3	2
145	1	1	2	3	2
146	2	1	2	3	2
147	3	1	2	3	2
148	1	2	2	3	2
149	2	2	2	3	2
150	3	2	2	3	2
151	1	3	2	3	2
152	2	3	2	3	2
153	3	3	2	3	2

154	1	1	3	3	2
155	2	1	3	3	2
156	3	1	3	3	2
157	1	2	3	3	2
158	2	2	3	3	2
159	3	2	3	3	2
160	1	3	3	3	2
161	2	3	3	3	2
162	3	3	3	3	2
163	1	1	1	1	3
164	2	1	1	1	3
165	3	1	1	1	3
166	1	2	1	1	3
167	2	2	1	1	3
168	3	2	1	1	3
169	1	3	1	1	3
170	2	3	1	1	3
171	3	3	1	1	3
172	1	1	2	1	3
173	2	1	2	1	3
174	3	1	2	1	3
175	1	2	2	1	3
176	2	2	2	1	3
177	3	2	2	1	3
178	1	3	2	1	3
179	2	3	2	1	3
180	3	3	2	1	3
181	1	1	3	1	3
182	2	1	3	1	3

183	3	1	3	1	3
184	1	2	3	1	3
185	2	2	3	1	3
186	3	2	3	1	3
187	1	3	3	1	3
188	2	3	3	1	3
189	3	3	3	1	3
190	1	1	1	2	3
191	2	1	1	2	3
192	3	1	1	2	3
193	1	2	1	2	3
194	2	2	1	2	3
195	3	2	1	2	3
196	1	3	1	2	3
197	2	3	1	2	3
198	3	3	1	2	3
199	1	1	2	2	3
200	2	1	2	2	3
201	3	1	2	2	3
202	1	2	2	2	3
203	2	2	2	2	3
204	3	2	2	2	3
205	1	3	2	2	3
206	2	3	2	2	3
207	3	3	2	2	3
208	1	1	3	2	3
209	2	1	3	2	3
210	3	1	3	2	3
211	1	2	3	2	3

212	2	2	3	2	3
213	3	2	3	2	3
214	1	3	3	2	3
215	2	3	3	2	3
216	3	3	3	2	3
217	1	1	1	3	3
218	2	1	1	3	3
219	3	1	1	3	3
220	1	2	1	3	3
221	2	2	1	3	3
222	3	2	1	3	3
223	1	3	1	3	3
224	2	3	1	3	3
225	3	3	1	3	3
226	1	1	2	3	3
227	2	1	2	3	3
228	3	1	2	3	3
229	1	2	2	3	3
230	2	2	2	3	3
231	3	2	2	3	3
232	1	3	2	3	3
233	2	3	2	3	3
234	3	3	2	3	3
235	1	1	3	3	3
236	2	1	3	3	3
237	3	1	3	3	3
238	1	2	3	3	3
239	2	2	3	3	3
240	3	2	3	3	3

241	1	3	3	3	3
242	2	3	3	3	3
243	3	3	3	3	3
244	1	1	1	1	4
245	2	1	1	1	4
246	3	1	1	1	4
247	1	2	1	1	4
248	2	2	1	1	4
249	3	2	1	1	4
250	1	3	1	1	4
251	2	3	1	1	4
252	3	3	1	1	4
253	1	1	2	1	4
254	2	1	2	1	4
255	3	1	2	1	4
256	1	2	2	1	4
257	2	2	2	1	4
258	3	2	2	1	4
259	1	3	2	1	4
260	2	3	2	1	4
261	3	3	2	1	4
262	1	1	3	1	4
263	2	1	3	1	4
264	3	1	3	1	4
265	1	2	3	1	4
266	2	2	3	1	4
267	3	2	3	1	4
268	1	3	3	1	4
269	2	3	3	1	4

270	3	3	3	1	4
271	1	1	1	2	4
272	2	1	1	2	4
273	3	1	1	2	4
274	1	2	1	2	4
275	2	2	1	2	4
276	3	2	1	2	4
277	1	3	1	2	4
278	2	3	1	2	4
279	3	3	1	2	4
280	1	1	2	2	4
281	2	1	2	2	4
282	3	1	2	2	4
283	1	2	2	2	4
284	2	2	2	2	4
285	3	2	2	2	4
286	1	3	2	2	4
287	2	3	2	2	4
288	3	3	2	2	4
289	1	1	3	2	4
290	2	1	3	2	4
291	3	1	3	2	4
292	1	2	3	2	4
293	2	2	3	2	4
294	3	2	3	2	4
295	1	3	3	2	4
296	2	3	3	2	4
297	3	3	3	2	4
298	1	1	1	3	4

299	2	1	1	3	4
300	3	1	1	3	4
301	1	2	1	3	4
302	2	2	1	3	4
303	3	2	1	3	4
304	1	3	1	3	4
305	2	3	1	3	4
306	3	3	1	3	4
307	1	1	2	3	4
308	2	1	2	3	4
309	3	1	2	3	4
310	1	2	2	3	4
311	2	2	2	3	4
312	3	2	2	3	4
313	1	3	2	3	4
314	2	3	2	3	4
315	3	3	2	3	4
316	1	1	3	3	4
317	2	1	3	3	4
318	3	1	3	3	4
319	1	2	3	3	4
320	2	2	3	3	4
321	3	2	3	3	4
322	1	3	3	3	4
323	2	3	3	3	4
324	3	3	3	3	4

Table 27: Possible profiles in non-binary choice sets.