Eliciting behavioural strategies in sequential decision problems

Jean Yves Jaffray, Antoine Nebout, Marc Willinger, . Gate. Groupe
d'Analyse Et de Théorie Economique

▶ To cite this version:

HAL Id: hal-02818406
https://hal.inrae.fr/hal-02818406

Eliciting behavioural strategies in sequential decision problems

Jean-Yves Jaffray, Antoine Nebout and Marc Willinger

May 2007

Preliminary version : not for quote

Abstract

We describe an experimental project on behaviour of individuals facing a sequential decision problem. Our purpose is to test whether experimental subjects satisfy the consequentialist hypothesis (Hammond, 1988) and to identify their behavioural strategies in a risky context Our experimental design allows us to distinguish behavioural strategies among non-consequentialist subjects. We distinguish subjects who apply Bayesian-Nash rationality in a multiple-selves game (Karni & Safra 1989) from those who apply a "cooperative strategy" among selves (Nielsen & Jaffray, 2006).

Keywords: Dynamic consistency, Consequentialism, Multiple "selves", time independence, Rank dependent utility, experimental design.

Two concepts, consequentialism and dynamic consistency, are central in dynamic decision theory. These concepts are both studied by philosophers and decision theorists. We first provide a brief survey of the early and recent literature on the topic, followed by a presentation of our experimental design and conclude with the practical problems we are facing in the set up of the experiment.

1. Literature Review

1.1. Seminal texts on consequentialism and dynamic consistency

The main difficulty of the dynamic consistency problem is to define a general formal framework in which it is possible to clearly define dynamic consistency. Hammond proposed the decision tree framework (Hammond, 1988) and two axioms : separabililty and consequentialism.

According to separability at each decision node of the decision tree, the decision maker's behaviour depends only of the "future", i.e. on the decision nodes that he still can reach and the associated payoffs. This axiom is an adaptation of the classical dynamic consistency axiom in decision theory to the decision tree framework.

According to the consequentialism axiom, the decision maker follows the same strategy in decisions trees with same normal form.

Moreover, Hammond assumes in his paper that all decisions trees and all probability distributions are available. Then he derives the reduction of compound lotteries as a consequence of the second axiom. The main result of Hammond is to show how preferences which violate the independence axiom may not only make behaviour dependent on the structure of the decision tree but also induce a strong form of inconsistency. His partial solution to the problem is to refine the space of consequences

Machina arrives at a similar conclusion (Machina, 1989). Most of the non-expected utility models induce time inconsistent behaviour and the properties of separability/non separability of the preferences must always be discussed with reference to a given state of consequences including a sufficiently refined description to incorporate relevant emotional states.

These two papers raise the incompatibility between the non-expected models and the dynamic consistency hypothesis
.
In his book on rationality and dynamic choice, McClennen develops the idea that, even if a decision maker does not satisfy the independence axiom, she can still avoid such irrationality trap as the money pump or the Dutch-book. He proposes two possible behaviours : sophisticated choice and resolute choice behaviour.

A *sophisticated* decision maker takes into account possible deviations that she might perform in the future and accordingly changes her choices in the present. In other words this strategy can be interpreted as the backward induction reasoning with perfect information and infinite temporal representation and thus does not tolerate any lack of rationality. A *resolute* decision maker chooses her best plan at the beginning of the dynamic problem and even though she might be willing to deviate from it at some point in the future, she nevertheless sticks to the original plan.

The discussion of McClennen stands on a normative perspective and proposes counter-arguments to the adoption of Hammond's consequentialist axioms as normative assumptions. We will now look at some papers which discuss possible rational behaviours even if they are not consequentialist.

1.2. Non consequentialism and strategies

The non-expected utility models, for example RDEU, do not assume the separability property of expected utility theory. According to the above mentioned literatur under reduction of compound lotteries and under the assumption that preferences at different decision nodes are identical (same utility function and same weighting function), preferences over the decision tree are not dynamically consistent. In particular, the sophisticated strategy, i.e., the strategy

generated by a standard rolling back of the decision tree, is likely to be stochastically dominated (Hammond, 1988).

Taking a rational RDEU decision Jaffray and Nielsen (2006) showed that dynamic consistency remains feasible, i.e. the decision maker can avoid dominated choices, by adopting a non-consequentialist behaviour, if her choices in a subtree possibly depend on what happens in the rest of the tree (Jaffray & Nielsen, 2006). Based on Jaffray's theorical framework (Jaffray, 1999) they define an algorithm which works by backward induction in the tree but differs from the standard dynamic programming algorithm.

This operational method stands on the resolute strategy first described by McClennen. The decision maker is associated with a "self" at each node of the decision tree and as these "selves" come from the same mind they are supposed to act in a cooperative with each other. A parameter of local preference flexibility is then defined to express the propensity of the selves to deviate at a node when the immediate gain is too important compared to the long term predefined outcome.

To preserve time consistency of decisions, the authors defined a weak hypothesis of rationality : "A decision maker is rational if his behavioural rule can never make him choose a first order stochastic dominated strategy". Thus at each decision node each self has to choose a particular action in a set of undominated strategies. This procedure is interesting as it can be implemented for all decision trees after the definition of particular utility and weighting function in the RDEU case.

In contrast, Karni and Safra (1989) suggest the idea of behavioural consistency in order to avoid inconsistency and maintain non-expected utility (Karni & Safra, 1989). They implement this notion by regarding the same decision maker at different decision nodes as different agents, and then taking the Bayesian — Nash equilibrium of this game. Such behavioural consistency assume a non-cooperative attitude of the successive selves. They apply their concept to a finite ascending bid auction game. When the utility functionals are both quasi-concave and quasi-convex, then there is an equilibrium in dominant strategies where each bidder continues to bid if and only if the prevailing price is smaller than his value.

These two papers rely on opposite hypotheses on the interaction of the "selves"[1] and they lead as we will show in our experimental design section to different predictions. However, they are both using a RDEU preference representation.

It has to be mentioned that the literature on the dynamic extension of other non-expected utility models is widely developed. On the subject, the paper of Klibanoff, Marinacci & Mukerji (2006) is probably the most accurate. More recent articles (Epstein & Schneider, 2003, Machina, Rustichini & Marinacci, 2005) try to reconcile ambiguity models with dynamic consistency.

1.3. Experimental literature

A paper by Hey & Paradisio (1999) presents an experiment which can be closely related to the one presented hereafter. It deals more with the attitude of the decision maker towards the late resolution of uncertainty than with non-consequentialist time consistency but it provides insights on the validity of the time independence axiom. We will describe this experiment in our presentation to compare and discuss the relevance with respect to ours.

2. Experimental design

In this section, for the sake of simplicity, we shall present the predictions of the two models in terms of EU. However, these predictions can be extended to the more general RDEU framework.

The experiment is based on the comparison of a subject's choices in three decision problems, which involve one-stage and two-stage lotteries. The outcomes of one-stage lotteries are monetary payoffs which will be denoted x_i. We shall assume $x_1 > x_1^* > x_2 > x_3$. The outcomes of two-stage lotteries are one-stage lotteries or final outcomes. Both types of lotteries will be denoted $(y, z ; q)$, where outcome y is obtained with probably q and outcome z with probability $(1 - q)$. The decisions problems involve two reference lotteries : $A = (x_1, x_3 ; q)$ and $B = (y, x_3 ; r)$ where y is a lottery.

---

[1] This expression is a bit misleading. It would be preferable to speak about brain areas or "imaginary/anticipated selves")

The three decision problems are depicted in figure 1 below, where circles correspond to chance nodes and squares to decision nodes (following the notations of Raiffa 1968). Problem 1 and problem 2 involve two chance nodes and one decision node. In problem 1 the decision node is in period 1 and the chance nodes in periods 2 and 3. In problem 2 there is a chance node in period 1 followed by a Problem 2 is constructed from problem 1 by interchanging the decision node and the first chance node and by replacing the x_1^* outcome by $x_1 > x_1^*$. In other words the delayed lottery in problem 1 is transformed into an immediate lottery.
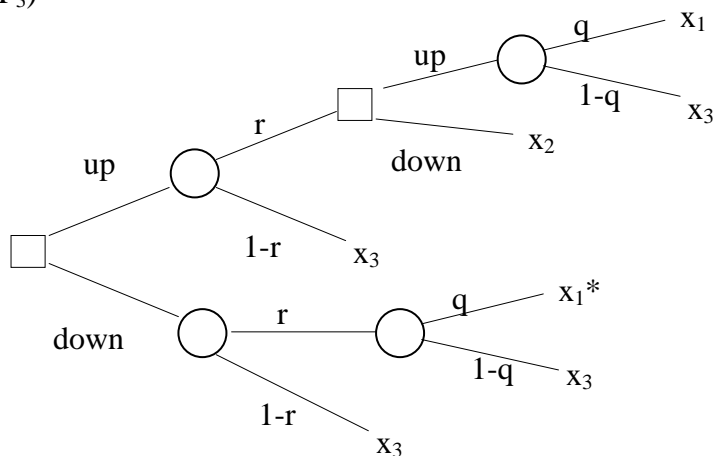
Problem 1 (P$_1$)
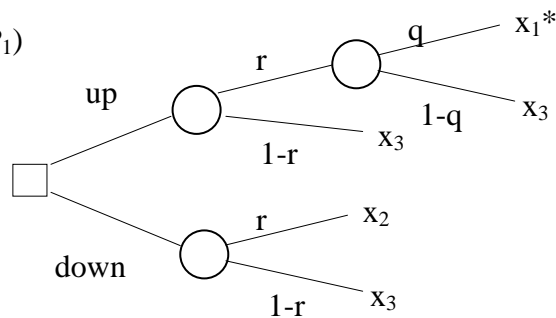
Problem 2 (P$_2$)
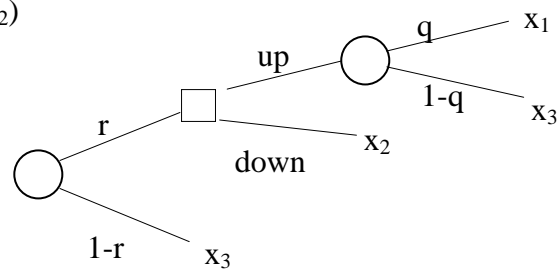
Problem 3 (P$_3$)

We first consider how a consequentialist decision maker would choose in problems P_1-P_3, and then we shall consider how a non consequentialist decision maker would choose in each case. Let us consider first, problems P_1 and P_2. We assume $x_1 > x_1^* > x_2 > x_3$.

A consequentialist decision maker, who chooses *up* in problem P_1, chooses also *up* in problem P_2. Similarly, a decision maker who chooses *down* in problem P_1 chooses also *down* in problem P_2. Consider an expected utility maximiser. Choosing *up* in P_1 has expected utility :

$$U_1^{up} = rqu(x_1^*) + [(1 - r) + r(1 - q)]u(x_3)$$

Similarly, by choosing *down* in P_1 the decision maker has expected utility :

$$U_1^{down} = ru(x_2) + (1 - r)u(x_3)$$

For problem P_2 the expected utility of choosing *up* and *down* are respectively :

$$U_2^{up} = rqu(x_1) + [(1 - r) + r(1 - q)]u(x_3)$$

$$U_2^{down} = ru(x_2) + (1 - r)u(x_3)$$

It is clear that since $U_1^{up} < U_2^{up}$ (by first order stochastic dominance), a decision maker who chooses up in problem 1 must also choose up in problem 2 and a decision maker who chooses down in problem 1 must also choose down in problem 2.

Let us consider now problem P_3. For an EU maximiser who chooses *up-up* the expected utility is equal to U_2^{up}. If he chooses *up-down* the expected utility is $U_1^{down} = U_2^{down}$. Choosing down leads to U_1^{up}. Therefore, a consequentialist decision-maker either chooses *up-up* in problem 3 or *up-down*. Specifically, he will choose *up-up* in problem 3 if he chose *up* in problem 1 and *up* in problem 2, and he will choose *up-down* in problem 3 if he chose *down* in problem 1 and *down* in problem 2.

Let us discuss now the decisions of a non-consequentialist decision-maker. First note that a non-consequentialist decision-maker either chooses *up* in problem 1 and *down* in problem 2, or *down* in problem 1 and *up* in problem 2. However, the choice in problem 3 will depend on the procedure implemented by the Selves.

Consider first the case where the decision maker chooses *down* in problem 1 and *up* in problem 2. If the current and future Selves play non-cooperatively (Karni & Safra, 1989), the current self will play *up* and the future self will play *up*. At the decision node 2, the future self

chooses $(x_1, x_3 ; q)$ over $(x_2 ; 1)$, since in problem 2 the decision maker revealed that he preferred $(x_1, x_3 ; q)$ when both options were available. At decision node 1 the current self therefore has to choose between $(x_1, x_3 ; q)$ which corresponds to playing *up,* and $(x_1*, x_3 ; q)$ which corresponds to playing *down*. By stochastic dominance the current self chooses *up*. If the current and future self behave cooperatively (Nielsen & Jaffray, 2006), they will either choose *up-up* or *up-down*. The future self selects an option set containing only undominated options. Clearly, playing *up,* i.e. choosing $(x_1, x_3 ; q)$, and playing *down*, i.e. choosing $(x_2 ; 1)$, are both undominated. For the current self, the choice set contains therefore $(x_1, x_3 ; q)$, $(x_2 ; 1)$ and $(x_1*, x_3 ; q)$, which corresponds to playing *down* in period 1. Since the latter option is dominated, the current self always decides to play *up*. Therefore, both the cooperative and the non-cooperative theory of the multiple selves predict that the current self will play up.

However, in the final case where the decision maker chooses *up* in problem 1 and *down* in problem 2, the two theories have opposite predictions. Again consider first how the non-cooperative selves will behave. The future self will choose down because he prefers $(x_2 ; 1)$, to $(x_1, x_3 ; q)$ according to his choice in problem 2. Therefore the choice set of the current self contains $(x_2, x_3 ; r)$ and $(x_1*, x_3 ; rq)$. Since the latter is preferred by the decision maker according to his choice in problem 1, the current self plays *down*. By a reasoning similar to the one used in the previous case, the cooperative present self will play up because playing down is dominated by playing up in the second period.

We can indeed summarize our predictions in this table

Hypothesis	P1-P2 \ P3	Up-Up	Up-Down	Down
Consequentialist	Up-Up	Yes	No	No
Consequentialist	Down-Down	No	Yes	No
Multiple Self	Up-Down	Nielsen-Jaffray	Nielsen-Jaffray	Karni-Safra
Mutiple Self	Down-Up	Nielsen-Jaffray Karni-Safra	Nielsen-Jaffray	No

So we can test for the validity of our hypothesis using a frequency approach. We denote freq (00/01) the number of subjects who have chosen 00 in P3 conditioning to the fact they choose 0 in P1 and 1 in P2.

To strengthen the consequentialist hypothesis we need :

Freq(Up-Up / Up-Up) > Freq(Up-Down + Down / Up-Up)

Freq(Up-Down / Down-Down) > Freq(Up-Up + Down / Down-Down)

To validate Nielsen-Jaffray hypothesis versus Karni-Safra one, we need:

Freq(Up-Up + Up-Down / Up-Down) > Freq(Down / Up-Down)

3. Pilot experiment and perspectives

3.1. Parameters

To run our experiment we have to set the probabilities r and q of our lotteries. Remind that in the above discussion we chose to present the predictions of the two models in terms of EU rather than the more general RDEU framework.

An important feature of our experiment is that the observations of choices in P2 and P3 depend of the realization of a lottery. In order to maximize our number of observations we decided to take r = 0.8 (in the Hey's experiment r = 0.2). To vary the probabilities between the two lotteries we set q = 0.7 in the second lottery[2]. Thus we need to evaluate if these values of probability are influencing our predictions. We could run this experiment for several probability distributions and see if the choice behaviour of the DM is significantly modified.

The second family of parameters we need to determine is the payoffs for each outcome of the lotteries. We decided for choose 1 < 6.5 < 9 < 10 Euros. In later experiments we will need to test the robustness of our results by choosing other probabilities and payoffs.

To control the influence of these parameters we intend to ask subjects to respond at the same type of questionnaire as in Holt & Laury (2002)at the beginning of a session.

---

[2] q = 0.8 in Hey's paper.

3.2. Organisation

To obtain independent observations from the subjects for each of the three problems, we intend to run the experiment at two different moments. First, we want them to answer P1 and P2, then run a totally different experiment and finally ask them to answer P3. In fact, it is important that the subjects do not remember their answers to P1 and P2 when they answer P3.

Subjects will be paid at the end by randomly choosing one of the three problems they answered to. We think that even if there can be an aggregation of probability bias, it is the best way to make them answer independently each problem. After one of the problems has been randomly selected we will run the lottery and use the answer of the subject to calculate his real gain.

3.3. Instructions

For now, we designed a paper questionnaire in which the probabilities are represented by the pulling of tickets of a ballot box. Since we intend to run a computer based experiment, we might add a visual presentation of the lotteries. Moreover to guarantee a good perception of the probabilities by the subjects we might include a training phase during which subjects will be accustomed to the outcomes of different lotteries.

We also think at adding a part at the end of the experiment. Before one of the three lotteries is randomly chosen we will determine which one of the three they will prefer to play. Using the same methodology than Hey and Paradisio (1999), we shall ask them to price the lottery by a second price auction or BDM's procedure. As Hey and Paradisio have already collected strong results for P1 and P2, we will focus on the difference that might exist between (P1/P2) and P3.

4. Conclusion

There is a growing theoretical literature on dynamic behaviour, but empirical evidence is still lacking on this topic. We propose a simple experiment based on lottery choice to gather new data about the way people are dealing with sequential decision problems.

One of our main concerns is to be certain that we are testing the right hypothesis. We need to be very cautious to avoid that our experiment induces subjects to make a decision plan ex-ante. If it is the case, we will fail to compare he behaviour of subjects facing different timing of resolution of uncertainty and consequently to test the dynamic consistency and consequentialism hypothesis. We will just have an insight at some counterfactual reasoning of subjects which is not our point here.. It is also worth mentioning that the test we propose is of course a within-subject test, allowing to test dynamic consistency for each subject separately. A between-subject test would be useless for our purpose.

5. Bibliography.

Epstein, L. (1980) "Decision Making and the Temporal Resolution of Uncertainty," *International Economic Review*, pp 267-281.

Hammond, P. (1988) "Consistent Plans, Consequentialism, and Expected Utility", *Econometrica*, vol.57(6), pp. 1445-1449.

Hey, J. and Paradiso, M. (1999) " Dynamic Choice and Timing-Independence: An Experimental Investigation" *Working paper*

Holt C, Laury S, (2002), "Risk Aversion and Incentive Effects in Lottery Choices" *American Economic Review*, vol 92, pp 1644-55.

Jaffray, J-Y. and Nielsen, T. (2006) "Dynamic decision making without expected utility : An operational approach", *European journal of operational research,* vol. 169(1), pp. 226-246.

Jaffray, J.-Y. (1999). "Rational decision making with imprecise probabilities".
 In 1st International Symposium on Imprecise Probabilities and Their Applications, pp. 183–188. Morgan Kaufmann Publishers.

Karni, E. and Safra, Z. (1989) "Dynamic Consitency, Revelations in Auctions and the Structure of Preferences". *Review of Economic Studies,* vol.56, pp. 421-434.

Klibanoff , P., Marinacci, M. and Mukerji, S. (2006) "Recursive Smooth Ambiguity Preferences", *Carlo Alberto Notebooks*

Kreps, D. and Porteus, E. (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, vol. 46, pp.185-200.

Machina, M. (1989). "Dynamic consistency and Non-expected Utility Models of Choice under Uncertainty", Journal of Economic Literature, vol.24(4), pp.1622-1668.

 McClennen, F. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: Cambridge University Press.

Rauchs A., Willinger M., (1996), "Experimental Evidence on the Irreversibility Effect". *Theory and Decision*, 40, 51-78.

Sarin, R. and Wakker, P. (1994) " Folding Back in Decision Tree Analysis", *Management Science,*vol.40(5), pp.625-628.

Savage,L.(1954). *The foundations of statistics*, Dover publications inc, New-York.

Segal, U. (1990), "Two-Stage Lotteries Without the Independence Axiom", *Econometrica,* vol. 58, pp. 349-77.