Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Economic Axioms in the Health Context

An Experiment about Kidney Transplant Preferences

Giacomo Battiston

BSc Thesis 2015-050

Laurea Triennale – Bachelor's Degree
in Economia Aziendale – Economics and
Management

Final Thesis

Economic Axioms in the Health Context: An Experiment about Kidney Transplant Preferences

Supervisor
Ch. Prof. Giacomo Pasini

Graduand
Giacomo Battiston
Matriculation Number 836845

Academic Year
2013 / 2014

1. Introduction

People's preferences differ with respect to a wide variety of things. People like to eat different things, they like to wear different clothes, they spend different amounts of time for the achievements they want to obtain, and so on. In general, markets are the means society use to respond to such diversity. However, on some occasion, market allocation is not available, due to market failures, or to ethical considerations. Allocation of cadaveric donors' kidneys among patients in waiting list for transplant is an example of an allocation problem that markets cannot solve. Trading organs for money is illegal and unethical[1] in every country of the world. However, since dialysis allows End-Stage Renal Disease (ESRD) patients to live without a functioning kidney, they could trade time in waiting list for organ quality. In order to know if such a system is viable, we should understand if patients' preferences differ. Discrete Choice Experiments (DCE) technique seems to be particularly well-equipped for eliciting preferences in the health domain, but we need to know that answers to DCE questionnaires are a trustable basis for analysis. DCEs often include a section testing the adherence of surveyed individuals to axioms of economic theory in the relevant field—otherwise only a maintained assumption. However, since DCE technique requires a careful experiment design in order to preserve statistical and response efficiency (Johnson et al., 2013), it is advisable to test whether including it in the experiment is necessary. This can be done by administering a questionnaire to a sample of individuals, before proposing the DCE to patients, as in this experiment.

DCE are a quantitative technique useful for eliciting preferences of individuals when revealed preferences cannot be used—for example in the context of goods that cannot be marketed or that have not been marketed yet. They are applied to goods or services with more than one attribute, where attributes can take a number of levels. In DCEs, individuals are administered a list of choice sets, including two or more alternatives, and asked to choose the preferred alternative for each. Choice alternatives are described by a fixed set of attributes, whose levels vary. Additionally, a status quo or opt-out alternative can be added for each choice set. Choice probabilities can be used to estimate the impact of a change in the level of an attribute on the probability of an

[1] Even though the existence of illegal organ markets suggests that for some groups of people in society this is not a taboo, it seems reasonable to say that the vast majority of people, across cultures, is opposed to it.

alternative being chosen, given some assumption about the latent utility function. This makes DCEs a useful tool in environments as diverse as marketing, transportation, and health (Ryan & Gerard, 2003). Also, the possibility to include continuous variables in the estimations performed from the experiments has been widely exploited, because of its obvious implications in marketing and policy. For example, this allows researchers to elicit willingness-to-pay (see, for example, Hole [2008], and Watson et al. [2004]), or willingness-to-wait (Watson et al., 2004; McIntosh & Ryan, 2002) for variations in attributes' levels.

As far as experiment design is concerned, DCEs involve a careful weighting of the questions to be included in the questionnaire. The number of choice sets proposed to the respondents has to be limited, in order to avoid fatigue effects. The selected choice sets should maximize statistical efficiency, usually represented by D-Efficiency (Johnson et al., 2013). Then, only the most informative questions should be integrated in the experiment. Inefficient questions, not to be integrated in the questionnaire, would normally include choice sets with a dominant alternative, or questions whose answer is already implied by transitivity. However, DCE literature documented the occurrence of violations both to non-satiation and transitivity axioms, and explored the phenomenon as an issue of internal validity (Louviere & Lancsar, 2009). For this reason, including some assumingly inefficient question in any DCE experiment in this context may be optimal. Literature might be useful in this sense, but violations may be context-specific to the particular choice proposed to the respondents.

In order to explore the issue, we performed a DCE on a sample of 101 second-year students of Ca' Foscari University of Venice. Students were proposed a number of questions in which they could violate non-satiation and transitivity axioms. For each question, respondents were asked to choose one of two grafts, defined on four dimensions. In this way, it was possible to perform three types of test: transitivity of preferences, non-satiation, and transitivity-and non-satiation[1]. Attributes used were realistic representations of the ones that could be used for a DCE of this kind, namely graft survival and waiting times, and infection and neoplastic risks[2] associated with the

[1] This was a test that could be failed by either violating non-satiation or violating transitivity of preferences.

[2] Infection and neoplastic risk are the risks that the transplanted organ infect the patient with a disease or bring a tumor inside their body, respectively.

graft. Choice sets and alternatives orders were randomized in order to control for order effects. The questionnaire's final section included three logic questions composing the Cognitive Reflection Test (CRT) (Frederick, 2005) and asked respondents to report their grades in the compulsory mathematics class of the course. This allowed for further analysis of preference axioms violations.

The experiment's results highlighted features of choice violations in the health domain. The most interesting result of the experiment was the high number of violations found. Each questionnaire allowed for a maximum of eight violations of axioms. 25.8% of the respondents violated an axiom in at least an occasion. Violations in non-satiation were found to be increasing with the ordinal position of the relative question. A test of this kind could be performed only on dominance questions, where only one question would be needed for detecting behavior contradicting economic theory. As for the covariates, only failing to answer correctly the first question of the CRT was found to be a significant predictor of the increase in the probability of violating axioms at least once, but the same test on each type of violation yielded slightly different results.

The results point toward two different mechanisms at work for the violations in non-satiation and transitivity. Specifically, violations in dominance and transitivity-and-non-satiation are significantly linked to errors in the CRT test, while violations in transitivity are not. This seem to suggest that, while violations in transitivity-and-non-satiation and non-satiation might be explained by low attention by a number of respondents, low involvement in the task, or cognitive abilities, this explanation is not completely justified by data in the case of violations in transitivity.

The next pages are organized as follows. First, theory and practice about DCE are briefly sketched. Then, a section is devoted to the theory of individual preferences, and its axioms. Subsequently, the approach of DCE to the testing of assumptions about individual preferences is examined. Next, attributes and levels to be used in the experiment are discussed. This provides guidance for the following section, where the design of the final questionnaire administered to the students is considered. Finally, the results of the experiment are analyzed and discussed and a concluding section deals with their implications for the development of DCEs in the field of kidney transplants.

2. DCE: Theory and Practice

DCE technique is grounded in Random Utility Theory (RUT), as formulated by Thurstone (1927) for choices involving pairs of alternatives, and extended by McFadden to sets of multiple alternatives (see, for example, McFadden [1986], and McFadden and Train [2000]), as explained in Louviere, Flynn, and Carson (2010).

RUT assumes that individual gain satisfaction, or "utility" for a choice alternative based on a latent utility function. This function is not observed, while choice probabilities in which it results are. Utility function is composed of an explainable part, dependent on attributes of the elements of the choice set and individuals' covariates. Also, it is composed of an unexplainable part, dependent on random factors. These are unidentified factors affecting choices of individuals. This concept is formalized in the following way (Louviere, Flynn, & Carson, 2010):

$$U_{in} = V_{in} + \varepsilon_{in}$$

Where U_{in} represents individual i's latent utility function for alternative i, V_{in} is the explainable part and ε_{in} is the unsystematic portion.

Given the random nature of ε_{in}, which alternative individual n will choose in the choice set C_n cannot be anticipated. Rather, based on some assumption about the distribution of ε_{in} leading to different models, we can estimate the probability that an individual n will choose the alternative i over any alternative j in the choice set he or she faces. This probability can be expressed as a function of the attributes of i and j and of individual's covariates, in the following way (Louviere, Flynn, & Carson, 2010):

$$P(i|C_n) = P[(V_{in} + \varepsilon_{in}) > Max(V_{jn} + \varepsilon_{jn})], \forall\, j \text{ in } C_n$$

The right assumption to be made about the distribution of the random component ε_{in} obviously depends on the type of choice explored by the experiment. While at first Thurstone assumed unobserved portions to be non-Independent-and-Identically-Distributed (non-IID) normal, an IID assumption had to be made in order to progress in the field, given that good approximate methods for solving integrals without closed forms were not available until recently (Louviere, Flynn, & Carson, 2010). McFadden

assumed an IID Gumbel-distributed random portion, a method yielding the closed form at the basis of the widely used Multinomial Logit Model (MNL) (Louviere, Flynn, & Carson, 2010). IID represents a strong assumption about the distribution of $\varepsilon_{in.}$ Specifically, given the following matrix represents a variance-covariance matrix for the unobserved portion of alternatives' utilities:

$$\begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix}$$

Assuming independence and identical distribution of the unobserved portion is tantamount to assuming the covariance terms (off-diagonal components) to be equal to 0, and all of the variances (on-diagonal components) to be equal across alternatives (Hensher, Rose, & Greene, 2005). This is represented by the following covariance matrix (Hensher, Rose, & Greene, 2005):

$$\begin{bmatrix} \sigma & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma \end{bmatrix}$$

As outlined by Hensher, Rose, and Greene (2005), making this assumption about the covariance matrix results into three straightforward weaknesses of the choice model. First, by assuming IID, it is necessarily implied that attributes not considered by the model have the same effect on each alternative. Second, a missing attribute common to two alternatives may introduce correlation between them, which is excluded by setting the off-diagonal terms to zero. Third, the inability to include the sources of variance into the observed part may result in different variance among alternatives, which is not allowed by IID.

Apart from IID, MNL assumes Independence of Irrelevant Alternatives, stating that "the ratio of choice probabilities are independent of the presence or absence of any other alternative in a choice set" (Hensher, Rose, & Greene, 2005). Even though this assumption may hold in many cases, in many others it results in mistaken predictions of choice probabilities (Train, 2002). Thus, IID and Indipendence-of-Irrelevant-Alternatives assumptions' limits, together with the fact that MNL does not allow for

preference heterogeneity, has prompted the exploration of new methods. On the one hand, closed-form-solution models could be obtained with a partial relaxation of assumptions (e.g. Nested Logit model). On the other hand, freer models were made possible by new methods of integration (Louviere, Flynn, & Carson, 2010), such as complete simulation (e.g. Probit model) or partial simulation (e.g. Mixed Logit model).

Apart from model-fitting considerations, DCEs also require a careful planning at the design stage. The problem of design is to find the combination of attribute levels describing each alternative in each choice sets yielding the most information about trade-offs, and resulting in unbiased estimates. Ideally, choice experiments should include all feasible choices (all the possible combinations of attribute levels). This would allow the researcher to estimate all main effects of each attribute and all interaction effects. However, it is clear that such designs would involve a huge number of questions, in most cases[1]. Then, it is normal to choose a "fractional design", selecting a manageable number of questions based on some criteria, such as the avoidance of implausible questions or questions with a dominant alternative, orthogonality, balance, and minimum overlap (Johnson et al., 2013). In order to have an efficient design, no choice set should be included for which the answer is already implied by previous knowledge about respondents' preferences. If transitivity-and-non-satiation apply to all of the respondents, choice sets including questions whose answers are implied by those axioms is inefficient. For similar reasons, implausible questions should not be asked[2]. Orthogonality means "independent variation of levels across attributes". Orthogonality is achieved if "each attribute level appears an equal number of times in combination with all other attribute levels". Balance requires that "all levels of each attribute appear an equal number of times" in the design. For a design to be balanced, it is necessary that the total number of alternatives is "divisible by the number of levels of each attribute" (p. 8). DCE design should also minimize overlap, that is the situation in which one or more attributes have the same level across alternatives of a choice set. It is true that this may reduce complexity and increase "response efficiency" (p.6); however, it also reduces the trade-off information given by answers, and then it lowers statistical

[1] For example, a choice experiment with 4 attributes, two with two levels, and two with four, would require [(2*2*4*4) * (2*2*4*4 – 1)]/2 = 2016 questions, for choice sets formed by two alternatives, to list all the possible combinations.
[2] These questions may also hamper respondents' trust in the experiment.

efficiency. Huber and Zwerina (1996) propose another criterion to be used in generating statistically efficient designs, utility balance. Utility balance occur when the expected probabilities of choosing the alternatives in a choice set approaches $1/N_c$, where N_c is the number of alternatives in a choice set. Then, utility balance criterion dictates to build alternatives for which utility is similar, in the same choice set.

"Manual" techniques for constructing efficient DCE designs are available (see for example Street, Burgess, Louviere [2005]). However, DCEs are often designed through specialized or non-specialized software, based on a measure D-Efficiency, in order to maximize statistical efficiency (Johnson et al., 2013). Maximizing D-Efficiency means finding a design "minimiz[ing] the joint confidence sphere around the complete set of estimated model parameters by maximizing the determinant of the inverse of the variance-covariance matrix in maximum-likelihood estimation" (p. 8).

In conclusion, DCE are a choice experiment tool, which requires the careful design of a choice questionnaire. Thus, testing assumptions about respondents' preferences is useful in order to avoid inefficient questions that may hinder the researcher's ability to fit the preference model.

3. Decision making under uncertainty: models and assumptions

An analysis of the issue of rationality in stated preferences can take different models of decision making as its starting point. In the special case dealt with by this thesis, respondents are decision makers who have to choose between options in a context of uncertainty. Respondents choose between lotteries in which they might assume a certain possible probability distribution over outcomes (or a number of them). Then, in the next few paragraphs, the main contributions to decision making in gamble choices are enumerated, in order to expound the theoretical framework employed by the thesis.

Von Neumann and Morgenstern (VNM) theory of decision making (1947) is one of the most influential contributions to microeconomic theory. Their work offered a framework for understanding choices through numerical utilities, setting the theoretical foundations for a quantitative analysis of individual's preferences. Their theory relies on four axioms of rational behavior: completeness, transitivity, and continuity of preferences, and independence of irrelevant alternatives. If, for utilities u and v, $u > v$ ($v > u$) indicates that u (v) is strongly preferred to v (u), and $u = v$ denotes indifference between u and v, completeness axiom states that for any utilities u, and v exactly one of the following holds:

$$u > v,\ v > u, \text{ or } u = v.$$

The axiom of transitivity affirms that for utilities u, v, and w*:*

$$\text{If } u > v \text{ and } v > w, \text{ then } u > w.$$

Continuity assumes that for utility values u, w, and v where $u < w < v$, there exists an α such that $0 < \alpha < 1$ and:

$$\alpha u + (1 - \alpha)v < w,$$

Also, there exists an α such that $0 < \alpha < 1$ and:

$$\alpha u + (1 - \alpha)v > w.$$

Finally, the axiom of independence of irrelevant alternatives[1] states that for u and v such that $v > u$, and for any w and for any α:

$$\alpha u + (1 - \alpha)w > \alpha v + (1 - \alpha)w.$$

The four axioms above define VNM rationality. This set of theoretic knowledge can then be applied to situations in which the decision maker's options are lotteries. These are probability distributions on a n of states of nature (outcomes). A lottery L is defined by:

$$L = \sum_{i=1}^{n} A_i p_i,$$

Where A_i is the i-th of n mutually exclusive outcomes, and p_i is the probability associated with that outcome, and:

$$p_1 + p_2 + \cdots + p_n = 1.$$

Given the axioms listed above and properties of the utility function $U(A)$, denoting a VNM-utility function attaching utility value u to outcome A, VNM proved that, for any VNM-rational agent:

$$L \prec M \text{ if and only if } E[u(L)] < E[u(M)],$$

$$\text{For } E[U(p_1 + p_2 + \cdots + p_n)] = p_1 U(A_1) + p_2 U(A_2) + \cdots + p_3 U(A_3),$$

The subsequent analysis of Savage (1954) expanded the analysis of VNM to the case in which the subject faces lotteries without knowing the probabilities attached to

[1] Interestingly, irrelevance of independent alternatives was not explicitly assumed by Von Neumann and Morgenstern; however, as explained by Malinvaud (1952), Samuelson proved that it was implied by their operations on equivalence classes. Specifically, their formulation of the operation implied Samuelson's "Strong Independence Axiom" (Samuelson, 1952).

each outcome with certainty. In this case, according to Savage and Ramsey the decision maker acts on her beliefs about the lotteries' probabilities (Ellsberg, 1961). Therefore VNM p_i is replaced by $P(x_i)$, representing the probability the subject associates with outcome x_i. Then, according to Savage, subjective beliefs about probabilities can be inferred from individuals' stated preferences (Ellsberg, 1961). For example, one may be asked to bet €50 on one of two players disputing a tennis match. If she chose to bet on tennis player A rather than B, she would reveal that she attaches to A no less probability of winning the match than to B, given the following properties of probabilities (Ellsberg, 1961):

I. there has to be a complete and transitive ordering of outcomes' probabilities;
II. if the outcome x_1 is more probable than x_2, then not-$x_1$is less probable than not-x_2;
III. if the pairs of outcomes x_1, x_2 and x_3, x_2 and x_1 is more probable than x_2, then the union $x_1 \cup x_2$ is more probable than the union $x_3 \cup x_2$.

Also, Savage proposes seven postulates about preferences (1954), three of which are addressed in the analysis of its work by Ellsberg (1961):

1. there has to exist a complete and ordering of lotteries for any given choice set;
2. choices between lotteries must be unaffected by the value of outcomes' pay-offs for which the lotteries payoffs are the same (sure-thing principle or independence of irrelevant alternatives);
3. dominated lotteries in a choice set are not chosen;

Ellsberg analyzed these postulates, and brought evidence against them. In particular, he proposed two paradoxical choice situations, for which individuals' preferences will likely violate the sure-thing principle.

In the first choice situation, the decision maker is presented with two urns. The first urn contains 100 black and reds balls, in unknown proportion. The second urn contains 50 black balls and 50 red balls. The decision maker is proposed four possible lotteries dependent on a one-ball random draw:

a) \$100 if red is drawn from the first urn;
b) \$100 if black is drawn from the first urn;
c) \$100 if red is drawn from the second urn;
d) \$100 if black is drawn from the second urn.

The decision maker would likely be indifferent between (a) and (b) and between the (c) and the (d). However, Ellsberg suggested that a number of individuals would prefer (c) and (d) to (a) and (b). This is a challenge to Savage's second condition about probabilities. Indeed, if the decision maker choices were restricted to betting on red, she would prefer to bet on the second urn, signaling that she deemed red in the second urn more probable than red in the first. Instead, if she was asked to bet only on blacks, she would still bet on the second urn, showing that she deems black in the second urn more probable than black in the second. Then, contrary to the second property shown above (II), her behavior would indicate that she regards "red in the second urn" more probable than "red in the first urn" and not-"red in the second urn" more probable than not-"red in the first urn". Ellsberg went on to demonstrate that such behavior is a violation of the first and the second Savage postulates as well, namely, the completeness postulate and the "sure-thing principle".

Another paradox proposed by Ellsberg (1961) casts doubt on Savage interpretation of gambles with uncertain probability distributions. As in the first examples, the decision maker is presented one urn, this time containing 30 red balls and 60 black and yellow balls, in unknown proportion. The decision maker is asked to choose one lottery between *I* and *II* and one between *III* and *IV*, described by the following tables, based on the outcome of a random draw from the urn.

	Red	*Black*	*Yellow*
I	\$100	\$0	\$0
II	\$0	\$100	\$0

	Red	*Black*	*Yellow*
III	\$100	\$0	\$100
IV	\$0	\$100	\$100

Ellsberg's prediction is that the decision maker proposed with such choice sets would choose the *I* lottery in the left-hand choice set and the *IV* lottery in the right-hand choice set. This behavior represents a clear violation of the "sure-thing principle", since, following Savage's second postulate, the change between the left-hand and the right-hand choice set should not influence which gamble is chosen. The only change between them is the payoff in the event "yellow", but this should not influence subject's choice, since its payoff remain equal between alternatives.

Ellsberg's interpretation of the behavior he described is that individuals are normally "conservative" (p. 667), that is averse to uncertainty in the definition of probabilities attached to lotteries. Ellsberg paradox and its interpretation of the effect of "ambiguity", or risk with undefined probability distributions resulted in models taking into account "uncertainties" as opposed to risks (see for example the "model of smooth ambiguity" of Klibanoff, Marinacci, Mukerji, [2005]). Also, it stimulated the development of other evidence against Savage's and Von Neumann and Morgenstern's contributions.

Work by Kahneman and Tversky (1979), too, took into question the assumptions of VNM expected utility theory. They proposed a new theory of choice of gambles, called "prospect theory", based on experimental evidence on subject's choices collected by them. Their hypothesis was that individuals' decision making occurs over two stages, editing and evaluation. Editing simplifies the representation of gambles (or "prospects", in Kanheman and Tversky words). In the stage of evaluation, instead the gambles are evaluated and the preferred gamble(s) is (are) chosen. Evaluation depends on two elements, $\pi(p)$, and $v(x)$. $\pi(p)$ is a function assigning decision weight of probabilities, and reflecting individuals' overweighting of small probabilities and underweighting of high ones. $v(x)$, instead, is a "value function", assigning value v to x, an outcome's deviation with respect to a reference point; this is a key difference with expected utility theory, where utility is defined by absolute, and not relative, wealth. $v(x)$ passes through the reference point; it is convex before the reference point and concave after. Also, it is asymmetric, in the sense that the negative value associated with losses is larger (in absolute terms) than the positive value associated with gains. Then, the decision maker displays risk-aversion in the gain domain, risk-seeking in the negative one, and loss-aversion.

The model put forward by Kanheman and Tversky was their proposed explanation for the experimental evidence they collected through a number of choice experiments. Loomes and Sugden advanced an alternative model for explaining the findings by Kanheman and Tversky, in a harsh-toned paper published in (1982). Their alternative model took into account the role of regret in decision making.

Loomes and Sugden claimed that, given a choice set of two gambles, an individual's ultimate utility from choosing a gamble does not depend only on the

consequence of the chosen gamble, given the occurred state of nature. It also depends on the consequence of the other gamble outcome for that state of nature. In simple terms, individuals normally experience regret if their choices result in the worst outcome for a given state of nature, and rejoicing when their choice is the best for the state of nature that occurs. Then, Loomes and Sugden defined two different types of utility functions. The *choiceless utility function* $C(.)$*,* unique up to an increasing linear transformation, attaches a real-valued utility index to each consequence x that the decision maker experiences without having chosen it. Instead, when a consequence is the outcome of a choice, it is valued with a different utility function. For a gamble chosen in a choice set of two, the *modified utility function* $M(.)$*,* given the occurrence of a state of nature, assigns a real-valued utility index to each pair of choiceless utilities (utility of the chosen lottery and utility of the unchosen one). Given x_1, the consequence of lottery L for state of nature s, and x_2, consequence of lottery M for state of nature s,

$$C(x_1) - M(x_1, x_2)$$

represents the regret (if negative) or rejoicing (if positive) from having chosen *L* over M. It is assumed that $M(.)$ is an increasing function of the choiceless utility of the chosen lottery, and a decreasing function of the unchosen one. Also, it is assumed that if the two choiceless utility are equal, the modified utility will be equal to them, too, therefore implying no regret or rejoicing.

Given that individuals experience regret and rejoicing, Loomes and Sugden assumed that they will take into account this two feelings when making decions, based on their expectations. For *n* possible states of the world, characterized by *n* probabilities (known or inferred by the individual), they defined action *A* as a lottery with a *n*-tuple of consequences. If the decision maker has to choose between action A_i and A_k, she will take into account her expectation of the expected modified utility generated by the two actions, over *n* possible states of nature. This expectation, for action A_i is:

$$E_i^k = \sum_{j=1}^{n} p_j m_{ij}^k.$$

The decision maker will prefer A_i if $E_i^k > E_k^i$, A_k if $E_k^i > E_i^k$, and she will be indifferent between the two actions if $E_i^k = E_k^i$. According to Loomes and Sugden theory, the expectation of choosing A_i over A_k is defined by:

$$E_i^k = \sum_{j=1}^{n} p_j[c_{ij} - c_{kj} + R(c_{ij} - c_{kj}) - R(c_{kj} - c_{ij})],$$

Where c_{ij} represents the choiceless utility from the consequence of A_i in state of nature j, and $R(.)$ represents a *regret-rejoice function,* "which assigns a real-valued index to every possible increment or decrement in choice utility" (p. 810). After having defined this expectations for choice sets of two actions, they showed that it was possible to expand the model to larger choice sets (p. 816).

Loomes and Sugden claimed that their theory was the most simple explanation for Kahneman and Tversky's experimental evidence, and the one to be preferred, brandishing Ockham's razor principle[1]. Most important to the aim of this thesis, they provided a theory of choice that relaxes transitivity as an assumption. Given their theory decision making under risk, Loomes and Sugden retained the "sure-thing" principle and dropped transitivity, contrary to what Kanheman and Tversky had done. In later work by Loomes and Sugden (1987), and by Loomes, Starmer, and Sugden (1991), violations of transitivity in regret theory were further explored from a theoretical perspective. Loomes, Stramer, and Sugden (1991) defined the types of preference cycles that were consistent with regret theory. The experimental evidence they gathered underscored that the most-frequently occurring violations of transitivity were consistent with regret theory (Loomes, Stramer, & Sugden 1991; Loomes & Taylor, 1992).

To be sure, other explanations have been proposed for intransitivity of preferences. One of the main contributions to this field is "Intransitivity of Preferences", by Tversky (1969). In it, the author lists many possible causes of intransitivity. Learning effects may occur when decision makers are asked to reveal their preferences. They may be compelled to think about their preferences in a field for the first time, and they may gradually learn what they prefer. Inconsistencies may occur even outside learning effects. Individuals' preferences fluctuate, and errors in the evaluative process are

[1] A problem-solving principle, stating that, among competing hypotheses, the one with the least assumptions is the one that should be selected.

possible. Additionally, individuals may use choice rules that yield intransitivity in pairwise choices. For example, consider the following example described in Tversky's paper. An individual has to choose in a pairwise fashion for the possible couples formed alternatives, *x, y,* and *z.* The three alternatives are described on a bi-dimensional basis, by attributes I and II. Their attributes' levels are reported in the following table:

	I	II
x	2ε	6ε
y	3ε	4ε
z	4ε	2ε

Suppose that the individual chooses her preferred alternative in each pairwise comparison in the following way: if the difference between the alternatives in attribute I is strictly grater than ε, choose the alternative with the highest level of attribute I. If the difference in the two levels is less than or equal to ε, choose the alternative with the best value of II. In this case, *x* is preferred in choice set (*x, y*), *y* is preferred in (*y, z*), and *z* is preferred in (*x, z*), therefore yielding an cycle of preferences. This is an example of how intransitivity may be engendered by simplifying rules of decision making.

Even though transitivity presents these problematic features, it has received scant attention in DCE literature, as we will see in the next session. However, given that DCEs usually assume axioms of the classical theory of decision making, it is important that these axioms be tested in such a framework.

4. Evaluation of Internal Validity in DCEs

Three reviews of DCE literature (Ryan & Gerard, 2003; De-Bekker Grob, Ryan, Gerard, 2012; Clark et al., 2014) have discussed, among other things, how DCE studies' authors deal with the issue of validity. Tests for external validity comparing stated preferences to actual behavior have not been widely performed, despite their usefulness, as reported by the three reviews. This is likely due to the fact that DCEs are most useful when revealed preferences (RP) data are absent, and this represents an obstacle to tests of external validity. Tests for internal validity, instead, are commonly included into DCEs, in different forms.

Internal validity tests encompass tests for theoretical validity (i.e. testing that the sign of parameters is as expected) (Ryan & Gerard, 2003; De-Bekker Grob, Ryan, Gerard, 2012), and tests for tenets of preference-based consumer theory (Lancsar & Louviere, 2006), defining some notion of rationality. In particular, tests for non-satiation, performed by proposing choice sets with dominated alternatives, have received much attention (Lancsar & Louviere, 2006). Tests for transitivity have been much less common in literature (Lancsar & Louviere, 2006), maybe because they are difficult to implement (Clark et al., 2014). This is rather problematic, since transitivity is a fundamental assumption for of the classic theory of preferences, while non-satiation only ensures "well-behaved indifference curves" (Lancsar & Louviere, 2006, pag. 801). Sen's contraction and expansion properties (Sen,1977; Sen, 1993) are widely tested in DCEs (Ryan & Gerard, 2003; De-Bekker Grob, Ryan, Gerard, 2012; Clark et al., 2014; Lancsar & Louviere, 2006). Sen's contraction property of preferences predicts that if A is chosen in choice set (A, B, C), A will be chosen if the set is reduced to (A, B). This is a test of regularity, in the sense that it predicts consistency of preferences[1] (Lancsar & Louviere, 2006). Sen's expansion property states that if A is chosen in (A, B), B will not be chosen in (A, B, C). Respondents' failure to employ compensatory decision making (lexicographic preferences) was also commonly tested in the past (Ryan & Gerard, 2003; De-Bekker Grob, Ryan, Gerard, 2012; Clark et al., 2014; Lancsar & Louviere, 2006), but the frequency of such test is now declining (Clark et al., 2014). However, tests for non-compensatory decision making present two main problems

[1] Curiously, contraction inconsistencies in Sen (1993) are dealt with as examples of inconsistencies of preferences that might be created by the researcher's imposition of preference rules.

(Lancsar & Louviere, 2006). The first is that fractional factorial designs' results might give the illusion of non-compensatory decision making, by failing to make more complex decision rules visible. The second problem is that the absence of instances of trading in the answers to a DCE questionnaire may depend on the lack of sufficient variation of attribute levels in the design. Thus, a wide variety of tests is available, but consequences of testing violations depend on the researcher.

Since internal consistency tests are generally included in the questionnaire to be used for preference elicitation, it is natural to ask what to do with the questionnaires where some definition of rationality has been violated. Leaving them out of the analysis can generate selection bias, since there is evidence that violations of rationality are linked to respondent characteristics; however, including them in the analysis might bias the parameter estimates (San Miguel, Ryan, & Amaya-Amaya, 2005). Louviere and Lancsar (2006) claim that preferences of such respondents should be used in the analysis. They maintain that deleting them might mean to exclude valid preferences, and that RUT is able to cope with some mistakes by respondents in revealing their preferences. Ryan and Bate (2001), instead, found that preferences of respondents violating non-satiation do not differ, and their evidence supports dropping respondents failing tests of internal consistency.

In conclusion, many techniques for testing "rationality" are available in literature. Also, those who want to test rationality during questionnaire should choose what to do with the data from questionnaires violating tests, in a way that balances selection bias and estimate bias.

5. Background on kidney allocation and selected attributes

A brief premise about the aim of the context of the experiment and the allocation system of kidneys in cadaveric transplants is useful for understanding the attributes chosen in the final design of the questionnaire.

The experiment's context is the opportunity to devise a system of organ allocation that incorporates recipients' preferences. The efficiency gain from incorporating recipients' choice in the allocation algorithm may be shown by a simple example. There are two hypothetical possible recipients: A and B. Possible recipient B is less risk averse than A, more prone to accept a lower-quality organ, and less patient, and they are perfectly equal in all other respects (age, health, HLA-typing, etc.). Time in waiting list is slightly higher for A. The present allocation system in the North of Italy would allocate the first arriving kidney to patient A, even though she might be willing to wait for a better organ. Then, maximum utility (average or minimum across patients) is not ensured[1]. Thus, it is possible to improve the system by having individual preferences recognized by the allocation algorithm. This can come in two ways. Either individuals express their preferences upon entering the waiting list, and these are held into account when allocating the organs, or individuals' preferences are inferred by the system through patients' covariates.

A theoretical paper by Zenios and Su (2006), and a Monte Carlo simulation by Zenios, Su, and Chertow (2004) developed a framework for understanding the issue of patient choice in kidney transplants from cadaveric donors in the USA. The first paper approaches the issue from the point of view of mechanism design. The authors examine a policy in which heterogeneous organs are allocated to n queues corresponding to n types of patients, where the candidate's type is only known to the candidate. Given different probabilities of arrival of kidneys and patients for different queues, patients join them based on expectations about others' actions, in order to maximize their utility—in terms of quality-adjusted life expectancy (QALE). The authors analyzed the outcome of two social welfare functions, one maximizing total utility (ensuring efficiency) and the other maximizing minimum utility (ensuring equity). They found that a system in which patients report their types in order to join the preferred queue

[1] This naturally depends on the size of patients A and B's utility parameters attached to the kidneys' quality, the expected time before the next kidney arrives, and the expected quality of the next organs.

achieves truth-telling among recipients and utility outcomes balancing equity and efficiency. The second paper, too, reports the impact on patients' quality-adjusted life years (QALY) of a system in which patients can choose among different levels of kidney quality (probability of graft survival), with heterogeneous probabilities of arrival. Recipients choose the minimum quality they are willing to accept, and their choice maximizes their QALY. The outcomes of this and other policies were Monte-Carlo simulated based on data provided by the United States Renal Data System. Interestingly, the authors found that incorporating recipient choice would increase QALY, decrease median waiting time, increase the likelihood of transplantation, and reduce the number of discarded kidneys with respect to the policy employed by the United Network of Organ Sharing (UNOS) and with respect to a policy used to allocate marginal kidneys to willing patients (UNOS/ECD). The results were strongly supported by sensitivity analyses performed by the researchers. The Monte Carlo simulation also confirmed the finding that the system balances efficiency with equity.

In the light of the findings of Zenios and Su (2006) and Zenios, Su, and Chertow (2004), the recognition of patients' preferences in kidney transplant appears to be a sensible option. As it was briefly sketched before, a stage preliminary to the proposal of a preference-based system in organ allocation is to test for heterogeneity in patients' attitudes towards graft attributes. This would inform policy-makers about the existence of an opportunity for intervention. The author's effort aims to offering support for a DCE in this field. In order to do this, some attributes of grafts should be selected for analysis, based on the present kidney allocation mechanism in the North of Italy.

The allocation of kidneys in transplants from cadaveric donors in Italy varies from place to place, depending on the location of the transplant center and of the health center where the organ is removed from the donor. There are three main geographic areas. Nord Italia Transplant Program (NITp), our area of interest, comprises much of the North of Italy. It includes the regions of Lombardy, Liguria, Veneto, Friuli-Venezia Giulia, Marche and the autonomous province of Trento.

In the NITp area, when the two kidneys are removed from a cadaveric donor, the first is assigned to the waiting list of the transplant center where nephrectomy was performed (local pool), and the second is made available to the whole NITp waiting list (general pool). In both cases, the allocation algorithm is articulated into four steps. In

the first step, organ-specific risk is evaluated. If the risk is low enough, meaning that the kidney does not present anatomical damages and there is an extremely low probability of infection or neoplasm transmission, the organ is made available for the whole list, depending on the following stages of the algorithm. In the second step, donor age determines if a kidney has to be used for special transplants, e.g. the pediatric program or urgent transplants, or if it has to be allocated to the ordinary waiting list in single kidney transplant (SKT). In the third step, the need for organs to be assigned to special programs is controlled, and they are given priority over the ordinary waiting list for SKTs. In the fourth step, organs are allocated to specific patients. Here, organs are allocated in a way prioritizing (i) sensitized patients (based on their PRA). Next, (ii) minimum donor-recipient HLA-mismatch is ensured. Then, (iii) precedence is given to patients with higher waiting time, by placing them in waiting time brackets (more than 10 years, 3-10 years, or 0-3 years). Last, (iv) age difference with the donor is minimized. Thus, the present system relies both on efficiency criteria (e.g. HLA-mismatch minimization, age difference minimization) and on an equity criterion (time in waiting list[1]).

Most importantly, the algorithm specifications suggest that factors maximizing graft survival are already held into account in the system. Also, the system takes into account risk levels at a preliminary stage. Additionally, it uses waiting time as a criterion for allocation. Then, it seems reasonable to ask patients' preferences about expected time of graft survival (as a function of the organ's attributes and the recipient's characteristics), risk generated by the organs other than graft failure, and waiting time. Further, it seems safe to assume that the patients perceive in different ways the risk of contracting infectious diseases and the risk of contracting a tumor.

Then, the final attributes chosen to be included in the experiment were infection risk, neoplastic risk, expected time of graft survival, and waiting time. The inclusion of four attributes in the experiment design was consistent with practice in literature (Ryan & Gerard, 2003; De-Bekker Grob, Ryan, Gerard, 2012; Clark et al., 2014).

As far as attributes' levels are concerned, for risk attributes, chosen levels were low and high. Chosen levels for the expected graft survival time attribute were 8, 10, 12,

[1] Time in waiting list would be an efficiency criterion, instead, if the algorithm gave precedence to patients with lower levels of it. This is because time in dialysis decreases significantly the probability of graft survival (Jadoul et al., 2012).

and 15 years. For the expected waiting time, instead, chosen levels were 0, 6, 12, and 36 months. Also, numeric attributes were expressed as average for a group, and minimum for graft survival and maximum of waiting time for another group, in order to detect the effect of this change on the likelihood of axiom violations. To the end of testing violations in dominance it was assumed that patients prefer lower levels of risk, lower levels of waiting time, and higher levels of expected graft survival. This seems a rather safe hypothesis that does not seem to impose the interviewer's preferences on patients, a problem shared by many contributions to DCE internal consistency literature (Lancsar, Louviere, 2006).

6. Design of the questionnaire

The questionnaire administered to students was designed so as to resemble the format followed by a DCE questionnaire to in the health context. Respondents were asked to imagine that they were patients of ESRD, and that they were in waiting list for transplant, and they were explained what this meant in terms of therapy and consequences on one's life[1]. Also, they were explained the meaning of each attribute[2]. This is a common procedure in health economics (see for example Chapman [1996], Porteous, Ryan, and Bond [2006], Scott, Watson, and Ross [2003]). This procedure was performed in two complementary ways. Students' questionnaire included a text with such instructions, but they were also explained those directions in advance, because this was considered to be less fatiguing. Each question, or choice set, asked the respondent to choose their favorite combinations of levels between two proposed. The number of questions chosen (16)[3] is slightly higher than usual practice in literature (Ryan & Gerard, 2003; De Bekker-Grob, Ryan, Gerard, 2012; Clark et al., 2014), but this allowed for the exploration of fatigue effects.

As it was briefly sketched before, two sets of questionnaires were created, one expressing graft survival and waiting times as averages, and the other expressing waiting time in terms of maximum time and graft survival time in terms of minimum time. This allowed to test the effect of such variation in format on the likelihood of occurrence of violations. The order of questions and the position of the alternatives were randomized, in order control for order and labeling effects. The three questions composing the Cognitive Reflection Test were included at the end, in order to avoid fatigue effects on the possible violations. By the same token, this may have reduced the respondents' ability to answer those questions; however, the aim of the choice was to safeguard the reliability of data on axioms' violations, as the focus of the research effort.

[1] In order to avoid fatigue effects, respondents were asked to imagine a simplified situation. Namely they were asked to imagine that they were undergoing hemodialysis, without considering peritoneal dialysis. This is because hemodialysis is much more common than peritoneal dialysis in the North of Italy (Ruggieri, 2014).

[2] A translation of the text proposed to the students is included in the appendix.

[3] The combination of the numbers of attributes levels (2, 2, 4, 4) and the number of questions reflects the possible choice of a DCE design aiming to achieve balance.

The final design included three distinct groups of questions. The first group was composed by six questions to test violations of transitivity ("Transitivity group" in the appendix). This group was composed by two subgroups of three questions each. Each subgroups followed a three-question format, with first choice set (x_A, x_B), second choice set (x_B, x_C), and third choice set (x_C, x_A), where x_A, x_B, and x_C are three alternatives composed of distinct combinations of attributes. A violation of transitivity would occur if a respondent reported $x_A P x_B$ and $x_B P x_C$ and $x_C P x_A$, or $x_B P x_A$ and $x_C P x_B$ and $x_A P x_C$. Then, violations of transitivity would arise in two of the eight possible combinations of answers for each subgroup.

The second group included questions two tests of violation of transitivity or non-satiation (Transitivity-and-non-satiation group" in the appendix). The group was subdivided into two subgroups of three questions. Four of the eight possible combinations of answers to the questions of a subgroup led to violations in preference axioms. In particular, each subgroup included an "anchor" choice set (x_A, x_B), and two other choice sets. As for the first subgroup's following two choice sets, one of them consisted of alternatives (x_C, x_A) with attribute-levels $x_C P x_B$, assuming non-satiation. The other choice set consisted of alternatives (x_D, x_B) with $x_D P x_A$, assuming non-satiation. Then, if the respondent chose x_B in first choice sets, non-satiation and transitivity of preferences would imply that x_C be chosen in the second question. If, otherwise, if x_A was chosen in the first question, transitivity-and-non-satiation would dictate that x_D be chosen in the third choice set. As far as the second subgroup's second and third questions are concerned, one was composed of alternatives (x_C, x_D) with attribute-levels $x_C P x_B$, assuming non-satiation. The other choice set consisted of alternatives (x_E, x_F) with $x_B P x_E$, and $x_F P x_A$ assuming non-satiation. Then, if the respondent chose x_B in first choice sets, non-satiation and transitivity of preferences would imply that x_C be chosen in the second question. If, otherwise, if x_A was chosen in the first question, transitivity-and-non-satiation would dictate that x_F be chosen in the third choice set.

Finally, four distinct questions were proposed in order to test violations in the non-satiation principle ("Non-satiation group", in the appendix). For each of the proposed choice sets (x_A, x_B), non-satiation implied the choice of x_B. Then, for each question, choosing x_A would mean to violate the non-satiation axiom.

As briefly sketched before, the questionnaire included the questions of the CRT, after the choice sets proposed, as translated in Italian by Piattelli Palmarini (2005). This aimed at controlling which of the two "ways of thinking" was employed by the respondent, between "system 1" and "system 2" (Frederick, 2005). "System 1" is a cognitive process relying on quick intuitions, automatically leading us through repetitive and mechanical tasks of our life. "System 2", instead, is a way of thinking involving reflection (e.g. complex algebraic calculations) (Stanovich & West, 2000). The "inventor" of CRT, Frederick (2005), claims that respondents pass this test if they are using "system 2", while they fail it when they use "system 1". Then, a link between this test and the probability to commit axioms is important in that it suggests which is the system individuals use in forming and revealing "rational" preferences. Importantly, the test has been found to be linked with IQ (higher IQ scores are linked to higher CRT scores), and sex (males' scores are higher than females' ones, on average) (Frederick, 2005). Additionally, higher CRT scores are linked to less risk aversion for gains and risk seeking for losses (Frederick, 2005). Then, a link between CRT results and violations would mean that dropping questionnaires violating rationality in DCEs may result in selection bias consequences.

In the final questionnaire, the order of the sixteen questions proposed was randomized, together with the order of the alternatives inside the questions. They were introduced by the instructions explained above, and CRT questions were asked next. The questionnaire administered to the students included a brief introduction, with an explanation of the attributes, 16 choice sets, the three CRT questions, and the mathematics' grade question[1].

[1] Two copies of the original questionnaire (one for each version) are included in Italian in the appendix, with a translation for all of its parts.

7. Results

The analysis of the data gathered through the experiment revealed a number of violations of the axioms tested, even though the vast majority of respondents did not violate any of the eight tests. Further, frequencies of single types of violations were low. Statistical tests performed underscored the non-random nature of these frequencies. Also, the occurrence of violations, in aggregate or separately, was not influenced significantly by the definition of times as average times as opposed to minimum or maximum times. Other covariates were found to have significant interactions with the probability of committing a violation, instead. CRT questions' answers were found to impact the probability of committing a violation significantly, though further analysis reveals that it has an impact only on the groups of questions involving dominance. The grade of the mathematics class had a more limited effect Additionally, some fatigue effects were found by the analysis, which can be relevant to implementations of DCEs experiments.

Before going through the results found, a cautionary note is necessary. 6 of the 101 questionnaires administered had one or more items non-responded. Specifically, one of them had one item non-responded out of 16, while the other had 15. The latter was not used in the analysis, while the others were. The specific questions that had not been answered by the five respondents were different questions (3, 8, 9, 12, and 16), and had different positions (1st, 2nd, 5th, 10th, and 11th). Then, it seems safe to assume that respondents involuntarily skipped these questions, and that they did not omit the answer for a particular characteristic of the question (such as complexity or position). In the analysis that follows, each of the numbers reported refers to the maximum sample available for the calculated the statistic, except for regressions, where the sample used was the maximum available for testing all the explored variables. This is because, by progressively dropping non-significant variables, a slightly lower sample might have been retained than the maximum obtainable. This has been done in order to preserve consistency and comparability among regressions.

Violations were found for each axiom. In general, 25.8% of respondent (N = 97) committed at least a violation of an axiom. 9.1% of the respondents (N = 99) reported an intransitive pattern of preferences. As far as the group of questions about transitivity-

Type of test	*Failure rate*	*N*	*σ*[1]	*p-value*[2]
Transitivity (1st group)	6.06	99	0.25	0.000
Transitivity (2nd gr.)	5.00	100	0.25	0.000
Transitivity/Non-satiation (1st gr.)	8.16	98	0.50	0.000
Transitivity/Non-satiation (2nd gr.)	5.05	99	0.50	0.000
Non-satiation (1st question)	8.00	100	0.50	0.000
Non-satiation (2nd q.)	5.00	100	0.50	0.000
Non-satiation (3rd q.)	2.00	100	0.50	0.000
Non-satiation (4th q.)	1.01	99	0.50	0.000
Transitivity (at least 1)	9.09	99	0.44	0.000
Transitivity/Non-satiation (at least 1)	12.24	98	0.44	0.000
Non-satiation (at least 1)	13.13	99	0.94	0.000
All types (at least 1)	25.77	97	0.98	0.000

[1] Probability of failing the test answering the questions randomly.
[2] For a one-tailed alternative hypothesis (failure rate < σ)

and-non-satiation are concerned, 12.1% committed a violation (N = 98). Concerning non-satiation, 13.1% chose a dominated alternative (N = 99).

Number of violations per questionnaire		
Number	*Percentage (%)*	*N*
1 violation	15.79	95
2 violations	5.26	95
3 or more	3.16	95

The frequencies found were compared to the probabilities of committing violations by filling the questionnaire by chance, so as to check for negligence. The probability of committing a violation answering by chance (σ), could be calculated for each subgroup of questions. This was 25% for each subgroup of questions about transitivity, and it was 50% for each group about transitivity-and-non-satiation and for each question about non-satiation". Through union of probabilities, it was found that the probability of committing at least a violation was 43.8% for transitivity, 75% for transitivity-and-non-satiation, 93.8% for non-satiation, and 98.0% for all violations. Then, a binomial test could be performed for each result found. Each of the listed probabilities was different from its counterpart generated by chance (subgroups, groups, and total).

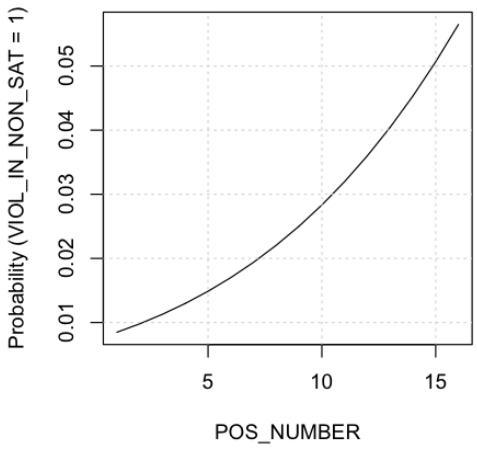
Chi-square tests were performed for detecting differences in frequencies of violations across subgroups. The goal was to find out if the probability of committing a violation in a subgroup (or in a single question, in the case of the non-satiation group) was different from the probability of committing a violation in another subgroup (or another single question). The only significant difference found was the one between the

Probit regressions						
	Estimate	*Std. Error*	*p-value*	*ME*[a]	*Error (ME)*	*N*
Dependent Variable						
atleast1viol[b]						92
Independent variable(s)						
Intercept	-12.096	0.2267	9.56e-08	-0.360	0.039	
crt1wrong[c]	10.483	0.3034	0.000551	0.323	0.091	
Dependent variable						
atleast1viol_tran[d]						92
Independent variable(s)						
Intercept	-26027	0.7328	0.000383	-0.001	0.343	
crt1wrong	12590	0.8644	0.145267	0.000	0.195	
maths22to26[e]	-17.4658	1796.5542	0.992243	-0.112	0.046	
Dependent variable						
atleast1viol_tran_non_sat[f]						92
Independent variable(s)						
Intercept	-2.8134	0.5944	2.21e-06	-0.254	0.038	
crt1wrong	1.4589	0.7145	0.0412	0.149	0.072	
Dependent variable						
atleast1viol_non_sat[g]						92
Independent variable(s)						
Intercept	-2.8134	0.5944	2.21e-06	-0.267	0.037	
crt1wrong	1.6094	0.7055	0.0225	0.174	0.075	
Dependent variable						
viol_in_non_sat						368
Independent variable(s)						
Intercept	-2.06473	0.42281	1.04e-06	-0.112	0.033	
pos_number[h]	0.05349	0.02977	0.0723	0.003	0.002	
crt1wrong	0.63359	0.28597	0.0267	0.039	0.020	
maths18to21[i]	-0.55823	0.32546	0.0863	-0.024	0.013	
maths22to26	-0.90169	0.37095	0.0151	-0.043	0.018	
maths28to30[j]	-0.76315	0.39700	0.0546	-0.029	0.013	

[a] Marginal effects calculated at the mean of the independent variables.
[b] Dummy variable taking value of 1 when at least a violation has been committed.
[c] Dummy variable taking value of 1 when the respondent answered incorrectly to the first CRT question.
[d] Dummy variable taking value of 1 when at least a violation of transitivity has been committed.
[e] Dummy variable taking the value of 1if the respondent's grade was between 22 and 26 out of 30.
[f] Dummy variable taking value of 1 when at least a violation of transitivity-or-non-satiation has been committed.
[g] Dummy variable taking value of 1 when at least a violation of non-satiation has been committed.
[h] Position of the question in the questionnaire.
[i] Dummy variable taking the value of 1if the respondent's grade was between 18 and 21 out of 30.
[j] Dummy variable taking the value of 1if the respondent's grade was between 27 and 30 out of 30.

frequencies of violations in the first and fourth questions of the non-satiation group. Frequencies of violations were respectively equal to 8% (N = 100), and 1% (N = 99).

By using the Probit model, the impacts of a number of covariates on the probabilities of violations were found as well. In particular, answering incorrectly the first answer of the CRT predicted a statistically significant increase in the probability of committing a violation of an economic axiom by 32.3%. Other questions' answers were found not to be significant predictors of the probability of a violation. This might have been due to the correlation between the probabilities of answering incorrectly different questions, but a significant relation of this kind existed only between the second and third questions' answers. As for transitivity, there was no significant predictor of the probability of a violation at the 5% level, but the Akaike's Information Criterion suggested to retain two covariates in the model[1]. One was the dummy variable informing of an incorrect answer to the first question of the CRT test. The second was the dummy variable signaling if the respondent had received a grade between 22 and 26 out of 30 for the mathematics class (both included). As for the group checking violations in transitivity-and-non-satiation, an incorrect answer to the first CRT question was the only significant predictor of an increase in the probability of a violation. This increase was of 14.9%. As far as violations in non-satiation are concerned, the probability of violating non-satiation increased by 17.4% if the first question of the CRT was answered incorrectly. This was the only significant predictor, in this case.



Marginal effect per position number.

Finally, for questions about dominance, it was possible to test if the position of the question the questionnaire increased the probability of answering in violation of the axiom. A small but quasi-significant effect was found. The probability of committing a

[1] There was, however, "quasi-significance" (p-value = 0.0589) for the dummy variable reporting incorrect answers to the first question of the CRT, if the variable reporting a mathematics' course grade between 22 and 26 was dropped, contrary to the suggestion of the AIC.

p-values for χ^2 test			
	Viol. in non-sat. (2nd q.)	*Viol. in non-sat. (3rd q.)*	*Viol. in non-sat. (4th q.)*
Violation in non-sataition (1st question)	0.5662	0.1048	0.04221[a]
Violation in non-sataition (2nd q.)		0.4416[a]	0.2183[a]
Violation in non-sataition (3rd q.)			1[a]
	Violation in non-sataition (3rd and 4th question)		
Violation in non-sataition (1st and 2nd question)	0.02872		
	Violation in non-sataition (2nd and 4th question)		
Violation in non-sataition (1st and 3rd question)	0.4741		
	Violation in non-sataition (2nd and 3rd question)		
Violation in non-sataition (1st and 4th question)	0.8064		
	Violation in transitivity (2nd group)		
Violation in transitivity (1st group)	0.9863		

[a] Warning message displayed by R software in calculating the statistic: "Chi-squared approximation may be incorrect."

violation increased by 0.5% for each unit increase in the position number of the question in the questionnaire. This supports the hypothesis of fatigue effects.

Symmetry was also tested, by checking if the order of alternatives in a dominance question had an effect on the probability of violating non-satiation in that question. Specifically, a dummy variable reporting violation of non-satiation in a question of dominance was regressed on the dummy variable signaling that the dominated alternative appeared on the left in the questionnaire. No significant effect was found. Another test of distance was performed for the subgroups concerning transitivity with dominance. It was tested if the distance between the two questions where the violation could occur had a significant effect on the probability of committing such violation. Even if these two subgroups were composed of three questions, the violation depended on the answer to the "anchor question", and to one of the remaining two depending on the answer to the anchor question (for a more precise explanation, see the section about experiment design). Here neither distance nor the square of it were found to have a significant impact on the probability of committing such a violation. The significance of distance might have accounted for a learning effect, for inconsistencies of preferences over time, or for fatigue effects. A similar test was performed on the subgroups of questions relating to violations of transitivity, to check the presence of any learning effect. A dummy variable reporting violations in transitivity, for the first and second subgroups, was regressed on the position number of the first question of the group to appear in the questionnaire. No significant relation was found.

8. Discussion

Results suggested that non-satiation is a safe assumption, while transitivity is slightly less so. Also, they gave insights about fatigue effects, they brought evidence against a dynamic process of the type described by San Miguel, Ryan, and Amaya-Amaya (2005), and against learning effects in general. Additionally, they indicated that defining times as minimum or maximum times, or average times does not seem to make a difference on the probability of violations.

A wrong answer to the first question of the CRT was a significant predictor of committing at least a violation in the questionnaire (32.3%). However, a closer look to the data reveals that a wrong answer to the first question of the CRT is a significant predictor only for violations of non-satiation (increase in probability of 17.1%) and of transitivity with dominance (increase in probability of 14.9%), while this is not the case for transitivity. This suggest a role for lack of involvement, lower cognitive abilities (Frederick, 2005), or use of "system 1" instead of "system 2" (Frederick, 2005) in increasing the probability of a violation in the first two categories. As it was briefly sketched above, males score higher on CRT, higher CRT scores are linked to higher IQ scores, and higher CRT scores are linked to higher risk-aversion for gains and risk-seeking for losses (Frederick, 2005). Then, a DCE excluding questionnaires of respondents who did not pass "rationality" tests may well translate into selection bias, and biased estimates.

Other CRT questions were less significant across regressions. This might be due to the correlation among them and the first, but a significant relationship of this kind was found only between the second and the third questions, using a Probit model. As far as violations in transitivity are concerned, data point toward an only partial effect of the same mechanism at work for the two previously dealt with. There seems to exist a difference with the other two types of violations, and to be sure, this would not be surprising, given that a variety of theoretical contributions have explained how "conscious" variations of transitivity are possible.

One of the theories of choice relaxing transitivity as an axiom, as explained before, is regret theory. The first theoretical paper about regret theory (Loomes, Sugden, 1982), and successive experimental contributions (Loomes, Stramer, & Sugden 1991; Loomes & Taylor, 1992) dealt with the issue, and set a theoretical framework for

understanding violations of transitivity. Most interestingly, Loomes, Stramer, and Sugden (1991), and Loomes and Taylor (1992) defined instances of preference cycles "predicted" under regret theory, in gambles in which the distribution of probabilities and payoffs are clear. For the questions of the present questionnaire, it is clearly impossible for the author to understand which distributions are assumed by the respondent, since all of the data are characterized by uncertainty. However, respondents, or a part of them may be using the decision-making model described by regret theory, and this does not ensure transitivity in their stated preferences.

Another possible explanation is given by the previously cited paper of Tversky (1969). Respondents, or a part of them, may have been using a decision rule which does not weight all the attributes into one function. As the previously exponded Tversky's example shows, they may have an algorithmic decision rule resulting in intransitive preferences. Think about the following example, based on the second subgroup in the transitivity group. The respondent may employ the following algorithm for choosing between two alternatives:

1. if the absolute difference in the expected graft survival is strictly larger than three years, choose the alternative with the higher expected graft survival time, if not, go to step 2;
2. if the difference between the expected waiting times in the two alternatives is different from zero, choose the alternative with the lower waiting time, if not, go to step 3;
3. if the neoplastic risk levels are not the same for the two alternatives, choose the alternative with the lower risk, if not, go to step 4;
4. if the infection risk levels are not the same for the two alternatives, choose the alternative with the lower risk level, if not, choose the one with higher expected graft survival time.

This choice rule would result in a cycle of the type $x_A P x_B$ and $x_B P x_C$ and $x_C P x_A$. Also, a choice rule of this kind does not seem implausible, given the complexity of the choice at issue. This problem could be particularly damaging for DCEs, since individuals choosing in this way would not be trading across attributes' levels in a traditional sense, and then research would not be able to infer preferred trade-offs.

Design, too, may have been partially responsible for the violations found for transitivity and transitivity-and-non-satiation. First, results might have depended, in part, on the lack of an indifference option ("Not sure"). In discussing their choice experiment, San Miguel, Ryan, and Amaya-Amaya (2005) suggest to explore the role of such issue on violations of rationality. If respondents, or a part of them, did not have clear preferences for one or another option, they probably would have chosen randomly, generating some of the inconsistencies. Second, there was no opt-out option ("Continue dialysis") for the questions. Respondents, all or part of them, might have been opposed to undergo a transplant for some of the choice sets. In particular, they might have profoundly disliked risk, so much so that, when one of the risk attributes was high, they saw the two proposed transplants as equally unacceptable.

Other effects, described by Tversky (1969), may be responsible for the results in transitivity violations. For example, simple intertemporal inconsistencies of preferences might be explaining the cycles. Also, learning could be advanced as an explanation. However, a rough test for learning was performed on data, and a significant effect was not found. As explained before, a Probit regression was performed where the dependent variable was a dummy variable signaling that the respondent had committed a violation in transitivity in the first or the second subgroup, while the independent variable was a variable reporting the position number of the first question of the relative subgroup in the questionnaire. The idea was that, if respondents were learning their preferences over the course of the questionnaire, subgroups with higher position number would have a lower probability of violations, because respondents had already gone through the other choice sets before coming to that[1]. No significant effect was found, for single or multivariate regressions.

As explained before, a similar test was performed on violations of transitivity-and-non-satiation. The dummy variable reporting violations of this category (for the first or the second subgroup in the group) was regressed on the distance between the first-occurring and the second-occurring question of the couple of questions where the violations could have happened. This might have signaled fatigue effects, learning effects or inconsistencies over time. Again, the variable's impact was not significant, in

[1] Here, it is obviously assumed that the respondents go through the choice sets in order. This assumption is supported by results about fatigue effects, which will be discussed later in the text.

single or multivariate regressions. Then, a process of learning does not seem to be at the basis of the results found for mixed transitivity-and-non-satiation.

Another type of dynamic process was tested on data about violations in transitivity with dominance. In Swait and Adamowicz (2001), it is proposed that respondents may be learning their preferences during the first questions, using the obtained knowledge about them to answer the next questions, and be fatigued in the last sets. San Miguel, Ryan, and Amaya-Amaya (2005) tested for this in DCE exploring violations of preference axioms. Specifically, they tested if the square of the distance between the position of the two choice sets within Sen's contraction and expansion tests was as significant predictor of the probability of passing those tests. San Miguel, Ryan, and Amaya-Amaya (2005) found a significant concave relationship between the dependent and the independent variables, and this was claimed to strengthen the claim by Swait and Adamowicz (2001). Thus, the same dynamic process was tested for violations in transitivity-and-non-satiation, by testing the impact of the square of the distance between the two choice sets. No significant effect was found.

To be sure, evidence of fatigue effects, though extremely weak, can be found in the results. If a dummy variable signaling if non-satiation has been violated in each question (N=368) is regressed on the available covariates, Akaike's Information Criterion suggest to retain the position of the choice set in the questionnaire as a variable, together with a variable signaling if the first question in CRT was wrong for the relative respondents, and dummy variables signaling if the relative respondent was in the group of mathematics' grades that went from 18 to 21, from 22 to 26, or form 27 to 30. The position number variable is "quasi-significant" (p-value=7.23%), and it has a (moderate) effect on increasing the probability of a violation (0.4%) for a unit increase in the position of the choice set in the questionnaire[1].

For non-satiation, it was also possible to test the assumption of that order of the alternatives did not matter in determining violations. The position of the alternative did not have a significant influence on the probability of violating non-satiation. This is consistent with the results of Ryan and Bate (2001), supporting symmetry.

Other factors seem to have had an impact on the probability of violating non-satiation in one of the four questions of the questionnaire. As explained above, chi-

[1] A more thorough description of the phenomenon is available in figure (page 28).

square tests were conducted to understand if there was a significantly different probability of passing a dominance test with respect to another. It was found that this was the case only between the first and the last dominance question in the DCE. In this case, frequency of violations was equal to 8% for the first question (N = 100), and 1% for the last question (N = 99). The other frequencies were different, too, but not significantly (the second question had a frequency of violations equal to 5% and the third 2%)[1]. Also, chi-square tests were performed to see if the distribution of violations between the aggregate of the first and the second questions (N = 200) and the aggregate of the third and the fourth (N = 199) were different. Frequencies of violations were statistically different between the two groups. Chi-square tests between the aggregate first and third questions (N = 200) and the aggregate second and fourth questions (N = 199), and between the aggregate first and fourth (N = 199) and the aggregate second and third (N = 200) did not yield significant differences. A number of explanations can be proposed for this. First, the last two questions' comparisons show variations in the levels of two attributes, while the first two display just one variation. Then, the variation in the last two might have been more clear to the respondents. This would confirm the finding of San Miguel, Ryan, and Amaya-Amaya (2005) that more differences in attribute levels decrease the probability of violating rationality tests. Second, variations in the first two alternatives are located in the last two attributes (displayed at the bottom of the tables displaying the choice set), while variations in the last choice sets include variations in then last attribute as well as in the first two (shown at the top of the tables describing the choice set, and then more visible). Thus, the vertical format of the alternatives' description may be responsible for the results. It is plausible that the more visible position of the first attributes made more difficult to commit a mistake in understanding the question. Third, risk attributes might be the key attributes among the proposed ones, and the complexity of the task may lead certain respondents to look only to those answers. Despite the variety of explanations that can be proposed for the issue at hand, most of them point toward the same direction; certain respondents likely tried to lower the effort associated with the questionnaire. They did so by using less time-or-

[1] Noticeably, non-satiation failure rates were similar to those found by San Miguel, Ryan, and Amaya-Amaya (2005), for the last two questions. However, they were much higher for the first two. Both groups had failure rates lower than the ones found in Ryan, Bate (2001).

energy-consuming choice rules than the "rational rule" of weighting all the attributes depending on one's preferences.

A radically different explanation takes into account the role of altruism in respondents' choices. Respondents, or a part of them, might have violated non-satiation because they stated that they would accept that someone else took the best organ in the choice set (see Cooper and Kagel [2013] for a review of experimental literature documenting the issue). This would be consistent with the frequencies of violations of non-satiation questions, in the following way; a higher number of violations were registered for choice sets for which the variation between the two alternatives was a one-level increase in one attribute. It is plausible that violations registered respondents' stated willingness to accept a lower-quality organ to increase the welfare of another, as long as the variation in quality was not too large. The age of respondents might have strengthened their altruistic behavior, because they possibly deemed themselves more capable to endure the effort of waiting more time in dialysis than younger or older individuals. This phenomenon may also be part of the explanation for violations in the mixed transitivity-and-non-satiation. Nevertheless, the questionnaire required respondents to indicate the "preferred" treatment. It was not asked by them to indicate the one they would choose in a real-world choice.

The issue of the opt-out alternative mentioned before may be an alternative explanation for the difference between frequencies of violations across questions of non-satiation. The first two questions did not offer variations in the level of risk, while the third an fourth did. Respondents may have not taken seriously questions offering high levels of risk, but they may have done so when it was possible to lower such level by choosing the dominant option.

Finally, part of the violations may be due to respondents' unwillingness to think carefully about such an unpleasant issue as kidney transplants. The questionnaire asked them to imagine that they were ESRD patients, but they, or a part of them, may simply be avoiding that. Nonetheless, students were not compelled to take the survey. If they really disliked the task, they would have likely refused to do that, instead of filling the questionnaire randomly.

In conclusion, results support several explanations. However, it is proposed here that factors generating violations in dominance are probably "mistakes" that could be

related to the use of "system 1" during the questionnaire[1]. Violations in transitivity, instead, might encompass more profound choice mechanisms as well, which should be investigated. Nonetheless, it is important to stress that respondents violating an axiom were a minority, and for each category they were less than 15%. This suggests this phenomenon can be dealt with through the error term of the random utility theory equation, as suggested in Ryan, Bate (2001), and in Ben Akiva (1985) as in Louviere and Lancsar (2006).

[1] This naturally raises the question of which is the system of thought patients would use if they had to make a real-world choice about their transplant.

9. Conclusions

This bachelor thesis has contributed to explore the issue of rationality in DCEs. A number of violations of rationality has been found. Many explanations have been proposed for understanding the issue. Further research could investigate the cause of violations, through a choice experiment followed by interviews with respondents. In the meanwhile, we can tentatively conclude that, even though violations of preferences' axioms have been registered, their number and quality does not seem to hamper the validity of DCEs as a research instrument. Also, if questions testing internal consistency are included in a DCE of this kind, dropping questionnaires of respondents failing them may bias results.

Cognitive abilities, involvement in the questionnaire task, or use of an intuition-based thinking system had a significant impact on violations of non-satiation and transitivity-and-non-satiation. This effect was not significant for violations in transitivity. This suggests that some other mechanism is at work for violations of transitivity. Several of them have been proposed. For example, respondents may be using a model of decision making similar to the one proposed by Loomes and Sugden (1982). Further exploration of this hypothesis is possible. If probability distributions over outcomes are not uncertain (as in the present experiment), Loomes, Stramer, and Sugden (1991), Loomes and Sugden (1987), and Loomes and Taylor (1992) offer a theoretical framework for understanding if violations of transitivity are consistent with regret theory. Another possible explanation is the use of simplified choice algorithm by patients, of the type described in Tversky (1969), transcending trade across attributes in a traditional sense. This can be investigated through a choice experiment followed by interviews with respondents. Also, the lack of an opt-out alternative, could have led some respondent to choose randomly a transplant in the proposed choice set, when they would have preferred to stay in dialysis if faced with such transplant choice set (this may have been a consequence of the level of risk associated with such transplants). Such appears to be supported by the different frequencies of violations in non-satiations. However, these differences may have other explanations, such as the employment of rules of thumb in decision making in order to lower effort, or "other-regarding preferences" (Cooper and Kagel, 2013).

Even though findings offer interesting instances of problems in carrying out DCEs, they do not seem to hinder DCEs' validity, given that frequencies of single types of violations are very low. However, they indicate that violations are linked to the "way of thinking" employed by respondents. This means that, if "control questions" were introduced in a DCE in the relevant field, and violations were found, dropping questionnaires violating axioms of classical theory of preferences could bias findings. This, together with the low frequencies of individual violations in the experiment, suggests that the number of these type of questions in the final DCE design should be kept extremely low.

Appendix

Instructions for the questionnaire

Imagine you are patient of a kidney disease, whose kidneys just stopped working, and that for this reason you have to undergo hemodialysis. Hemodialysis is a treatment substituting kidneys and it consists of the filtering of the patient's blood through dedicated instruments. The patient is connected to the filtering machine through a arteriovenous fistula created for the occasion. Hemodialysis is usually performed in sessions lasting about 4 hours, 3 times per week, in a specialized medical center. This therapy has some noteworthy side effects, such as sense of fatigue and hypotension.

Imagine that you enrolled in waiting list for a kidney transplant. A treatment (transplant) is described along the following factors:

• infection risk (low or high) is the risk of contracting an infectious disease through the graft. If it is low, the organ the organ has undergone all the possible checks, even if complete safety cannot be guaranteed. If it is high, some of the controls have not been performed, or the donor had some risky behaviors in the days before his or her death, but an infection may still not result from checks (even if it is possible);

• neoplastic risk (low or high) is the risk of contracting a tumor through the transplanted organ. If it was low, the donor was not affected by a tumor, almost surely, even if a minimum level of risk does exist (for example if the donor was not aware of the problem and it did not emerge from checks). It is high if the donor had some kinds of neoplastic disease. Still, it is not high in terms of probability, because the due checks have been performed.

• graft survival is determined by the characteristics of the transplanted graft, by the characteristics of the recipient, an by compatibility between donor and recipient. It is stated in terms of minimum[1] of graft survival years.

• waiting time is the time one will have to wait in order to obtain the transplant proposed. It is stated in maximum number of years.

Below are proposed 16 pairs of treatments (transplants) described by different attributes. Please, indicate the preferred one for each pair, by crossing (X) the square below it.

[1] "Minimum" and "maximum times" were "expected times" in the other version of the questionnaire.

Questionnaire

Transitivity group

1st subgroup

Treatment A	**Treatment B**
Infection Risk: low	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: low
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 6 months	Maximum waiting time: 36 months
☐	☐

Treatment A	**Treatment B**
Infection Risk: low	Infection Risk: high
Neoplastic Risk: low	Neoplastic Risk: low
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 12 years
Maximum waiting time: 36 months	Maximum waiting time: 36 months
☐	☐

Treatment A	**Treatment B**
Infection Risk: high	Infection Risk: low
Neoplastic Risk: low	Neoplastic Risk: high
Minimum survival time of the graft: 12 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 36 months	Maximum waiting time: 6 months
☐	☐

2nd subgroup

Treatment A	**Treatment B**
Infection Risk: low	Infection Risk: high
Neoplastic Risk: low	Neoplastic Risk: low
Minimum survival time of the graft: 12 years	Minimum survival time of the graft: 15 years
Maximum waiting time: 6 months	Maximum waiting time: 6 months
☐	☐

Treatment A	**Treatment B**
Infection Risk: high	Infection Risk: low
Neoplastic Risk: low	Neoplastic Risk: low
Minimum survival time of the graft: 15 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 6 months	Maximum waiting time: 0 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: low	Neoplastic Risk: low
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 12 years
Maximum waiting time: 0 months	Maximum waiting time: 6 months
☐	☐

Transitivity-and non-satiation group

3rd subgroup

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: low
Minimum survival time of the graft: 8 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 0 months	Maximum waiting time: 12 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: low	Neoplastic Risk: high
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 8 years
Maximum waiting time: 12 months	Maximum waiting time: 0 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: low
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 8 years
Maximum waiting time: 0 months	Maximum waiting time: 12 months
☐	☐

4th subgroup

Treatment A	Treatment B
Infection Risk: high	Infection Risk: low
Neoplastic Risk: low	Neoplastic Risk: high
Minimum survival time of the graft: 15 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 36 months	Maximum waiting time: 0 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: high
Neoplastic Risk: high	Neoplastic Risk: high
Minimum survival time of the graft: 12 years	Minimum survival time of the graft: 15 years
Maximum waiting time: 0 months	Maximum waiting time: 36 months
☐	☐

Treatment A	Treatment B
Infection Risk: high	Infection Risk: high
Neoplastic Risk: high	Neoplastic Risk: low
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 15 years
Maximum waiting time: 0 months	Maximum waiting time: 12 months
☐	☐

Non-satiation group

Treatment A	Treatment B
Infection Risk: high	Infection Risk: high
Neoplastic Risk: low	Neoplastic Risk: low
Minimum survival time of the graft: 12 years	Minimum survival time of the graft: 15 years
Maximum waiting time: 0 months	Maximum waiting time: 0 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: high
Minimum survival time of the graft: 10 years	Minimum survival time of the graft: 10 years
Maximum waiting time: 36 months	Maximum waiting time: 12 months
☐	☐

Treatment A	Treatment B
Infection Risk: high	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: high
Minimum survival time of the graft: 15 years	Minimum survival time of the graft: 15 years
Maximum waiting time: 6 months	Maximum waiting time: 0 months
☐	☐

Treatment A	Treatment B
Infection Risk: low	Infection Risk: low
Neoplastic Risk: high	Neoplastic Risk: low
Minimum survival time of the graft: 12 years	Minimum survival time of the graft: 12 years
Maximum waiting time: 36 months	Maximum waiting time: 12 months
☐	☐

Questionario:

Immagini di essere un paziente affetto da una malattia renale, che ha appena perso l'utilizzo di entrambi i reni e che come tale deve sottoporsi a trattamento emodialitico. L'emodialisi è un trattamento che sostituisce le funzioni renali e consiste nel filtraggio del sangue del paziente attraverso apposita strumentazione. Il paziente viene connesso al filtro attraverso una fistola arterio-venosa opportunamente allestita. L'emodialisi viene generalmente svolta in sedute di circa 4 ore, 3 volte a settimana, in un centro medico specializzato. Questa terapia presenta delle notevoli controindicazioni, come stanchezza e ipotensione.

Immagini di essersi iscritto in lista d'attesa per ottenere un trapianto di rene. Un trattamento (trapianto) è caratterizzato dai seguenti fattori:

- il rischio infettivo (basso o alto) è il rischio di contrarre una malattia infettiva attraverso l'organo trapiantato. Se è basso, l'organo trapiantato è stato sottoposto a tutti i controlli necessari, anche se la totale sicurezza di esso non può essere garantita. Se il rischio è alto, alcuni controlli non sono stati effettuati, oppure il donatore aveva tenuto dei comportamenti a rischio infettivo nei giorni precedenti alla sua morte ma un'eventuale infezione potrebbe ancora non risultare dai controlli, pur essendo possibile;
- il rischio neoplastico (basso o alto) è il rischio di contrarre un tumore attraverso l'organo trapiantato. Se è basso, il donatore quasi certamente non era affetto da tumori che potessero interessare il rene, anche se un minimo livello di rischio esiste (ad esempio se il donatore non ne era al corrente e la patologia non è emersa dai controlli). È alto nel caso il donatore fosse affetto da particolari tipi di patologia tumorale. Rimane comunque un rischio non elevato in termini di probabilità, perché i dovuti controlli vengono effettuati;
- la sopravvivenza dell'organo trapiantato è determinata dalle caratteristiche dell'organo trapiantato, dalle caratteristiche del ricevente e dalla compatibilità tra donatore e ricevente. È espressa in numero minimo d'anni di sopravvivenza dell'organo;
- il tempo d'attesa è il tempo che si dovrà attendere per ottenere il trapianto proposto nell'alternativa. È espresso in numero massimo di mesi d'attesa.

Di seguito sono proposte 16 coppie di trattamenti (trapianti) descritti da attributi diversi. Per favore indichi il preferito in ciascuna delle 16 coppie, ponendo una X nel riquadro sotto di esso.

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 12 anni	Sopravvivenza minima dell'organo trapiantato: 12 anni
Tempo d'attesa massimo: 36 mesi	Tempo d'attesa massimo: 12 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 10 anni
Tempo d'attesa massimo: 6 mesi	Tempo d'attesa massimo: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 8 anni	Sopravvivenza minima dell'organo trapiantato: 8 anni
Tempo d'attesa massimo: 12 mesi	Tempo d'attesa massimo: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 8 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 12 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 10 anni
Tempo d'attesa massimo: 12 mesi	Tempo d'attesa massimo: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 15 anni	Sopravvivenza minima dell'organo trapiantato: 10 anni
Tempo d'attesa massimo: 36 mesi	Tempo d'attesa massimo: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 15 anni	Sopravvivenza minima dell'organo trapiantato: 15 anni
Tempo d'attesa massimo: 6 mesi	Tempo d'attesa massimo: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: alto
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 15 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 12 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 12 anni	Sopravvivenza minima dell'organo trapiantato: 10 anni
Tempo d'attesa massimo: 36 mesi	Tempo d'attesa massimo: 6 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: alto	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 15 anni	Sopravvivenza minima dell'organo trapiantato: 12 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 0 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 12 anni
Tempo d'attesa massimo: 36 mesi	Tempo d'attesa massimo: 36 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 15 anni	Sopravvivenza minima dell'organo trapiantato: 12 anni
Tempo d'attesa massimo: 6 mesi	Tempo d'attesa massimo: 6 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 12 anni	Sopravvivenza minima dell'organo trapiantato: 15 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 36 mesi
☐	☐

Trattamento A	Trattamento B
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 8 anni
Tempo d'attesa massimo: 12 mesi	Tempo d'attesa massimo: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 15 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 6 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza minima dell'organo trapiantato: 10 anni	Sopravvivenza minima dell'organo trapiantato: 12 anni
Tempo d'attesa massimo: 0 mesi	Tempo d'attesa massimo: 6 mesi
☐	☐

Per le seguenti domande inserisca la sua risposta nell'apposito spazio.

Una racchetta e una pallina da ping-pong costano, insieme, 1,10 €. La racchetta costa 1,00 € in più della pallina. Quanto costa la pallina?

_____

Ci sono 5 macchine che producono 5 prodotti finiti in 5 minuti. Quanto tempo occorrerà a 100 di quelle macchine per produrre 100 di quei prodotti finiti?

_____

In un lago ci sono delle ninfee che raddoppiano la loro superficie ogni giorno. Le ninfee impiegheranno 48 giorni a coprire tutto il lago. Quanti giorni impiegheranno a coprire metà del lago?

_____

Per favore, inserisca di seguito la votazione del suo esame di matematica. Nel caso in cui non avesse sostenuto l'esame ponga una X nello spazio indicato. Se invece non volesse dare quest'informazione, lasci pure lo spazio vuoto.
Votazione: __

Grazie per la sua collaborazione.

Questionario:

Immagini di essere un paziente affetto da una malattia renale, che ha appena perso l'utilizzo di entrambi i reni e che come tale deve sottoporsi a trattamento emodialitico. L'emodialisi è un trattamento che sostituisce le funzioni renali e consiste nel filtraggio del sangue del paziente attraverso apposita strumentazione. Il paziente viene connesso al filtro attraverso una fistola arterio-venosa opportunamente allestita. L'emodialisi viene generalmente svolta in sedute di circa 4 ore, 3 volte a settimana, in un centro medico specializzato. Questa terapia presenta delle notevoli controindicazioni, come stanchezza e ipotensione.

Immagini di essersi iscritto in lista d'attesa per ottenere un trapianto di rene. Un trattamento (trapianto) è caratterizzato dai seguenti fattori:

- il rischio infettivo (basso o alto) è il rischio di contrarre una malattia infettiva attraverso l'organo trapiantato. Se è basso, l'organo trapiantato è stato sottoposto a tutti i controlli necessari, anche se la totale sicurezza di esso non può essere garantita. Se il rischio è alto, alcuni controlli non sono stati effettuati, oppure il donatore aveva tenuto dei comportamenti a rischio infettivo nei giorni precedenti alla sua morte ma un'eventuale infezione potrebbe non ancora risultare dai controlli, pur essendo possibile;
- il rischio neoplastico (basso o alto) è il rischio di contrarre un tumore attraverso l'organo trapiantato. Se è basso, il donatore quasi certamente non era affetto da tumori che potessero interessare il rene, anche se un minimo livello di rischio esiste (ad esempio se il donatore non ne era al corrente e la patologia non è emersa dai controlli). È alto nel caso il donatore fosse affetto da particolari tipi di patologia tumorale. Rimane comunque un rischio non elevato in termini di probabilità, perché i dovuti controlli vengono effettuati;
- la sopravvivenza dell'organo trapiantato è determinata dalle caratteristiche dell'organo trapiantato, dalle caratteristiche del ricevente e dalla compatibilità tra donatore e ricevente. È espressa in numero medio d'anni di sopravvivenza dell'organo;
- il tempo d'attesa è il tempo che si dovrà attendere per ottenere il trapianto proposto nell'alternativa. È espresso in numero medio di mesi d'attesa.

Di seguito sono proposte 16 coppie di trattamenti (trapianti) descritti da attributi diversi. Per favore indichi il preferito in ciascuna delle 16 coppie, ponendo una X nel riquadro sotto di esso.

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 12 anni	Sopravvivenza media dell'organo trapiantato: 12 anni
Tempo d'attesa medio: 36 mesi	Tempo d'attesa medio: 12 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 10 anni
Tempo d'attesa medio: 6 mesi	Tempo d'attesa medio: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 8 anni	Sopravvivenza media dell'organo trapiantato: 8 anni
Tempo d'attesa medio: 12 mesi	Tempo d'attesa medio: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 8 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 12 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 10 anni
Tempo d'attesa medio: 12 mesi	Tempo d'attesa medio: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 15 anni	Sopravvivenza media dell'organo trapiantato: 10 anni
Tempo d'attesa medio: 36 mesi	Tempo d'attesa medio: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 15 anni	Sopravvivenza media dell'organo trapiantato: 15 anni
Tempo d'attesa medio: 6 mesi	Tempo d'attesa medio: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: alto
Rischio neoplastico: alto	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 15 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 12 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 12 anni	Sopravvivenza media dell'organo trapiantato: 10 anni
Tempo d'attesa medio: 36 mesi	Tempo d'attesa medio: 6 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 15 anni	Sopravvivenza media dell'organo trapiantato: 12 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 12 anni
Tempo d'attesa medio: 36 mesi	Tempo d'attesa medio: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: alto	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 15 anni	Sopravvivenza media dell'organo trapiantato: 12 anni
Tempo d'attesa medio: 6 mesi	Tempo d'attesa medio: 6 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: alto	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 12 anni	Sopravvivenza media dell'organo trapiantato: 15 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 36 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: alto
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 8 anni
Tempo d'attesa medio: 12 mesi	Tempo d'attesa medio: 0 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: alto
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 15 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 6 mesi
☐	☐

Trattamento A	**Trattamento B**
Rischio infettivo: basso	Rischio infettivo: basso
Rischio neoplastico: basso	Rischio neoplastico: basso
Sopravvivenza media dell'organo trapiantato: 10 anni	Sopravvivenza media dell'organo trapiantato: 12 anni
Tempo d'attesa medio: 0 mesi	Tempo d'attesa medio: 6 mesi
☐	☐

Per le seguenti domande inserisca la sua risposta nell'apposito spazio.

Una racchetta e una pallina da ping-pong costano, insieme, 1,10 €. La racchetta costa 1,00 € in più della pallina. Quanto costa la pallina?

_____

Ci sono 5 macchine che producono 5 prodotti finiti in 5 minuti. Quanto tempo occorrerà a 100 di quelle macchine per produrre 100 di quei prodotti finiti?

_____

In un lago ci sono delle ninfee che raddoppiano la loro superficie ogni giorno. Le ninfee impiegheranno 48 giorni a coprire tutto il lago. Quanti giorni impiegheranno a coprire metà del lago?

_____

Per favore, inserisca di seguito la votazione del suo esame di matematica. Nel caso in cui non avesse sostenuto l'esame ponga una X nello spazio indicato. Se invece non volesse dare quest'informazione, lasci pure lo spazio vuoto.
Votazione: __

Grazie per la sua collaborazione.

References

Ben-Akiva, M. E., & Lerman S. (1985). *Discrete choice analysis: Theory and applications to travel demand*. The MIT Press: Cambridge.

Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(3), 771-791.

Cooper, D. J., & Kagel, J. H. (2013). Other-regarding preferences: A selective survey of experimental results. *Handbook of Experimental Economics, Volume 2.* Princeton: Princeton University Press.

De Bekker-Grob, E., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: A review of the literature. *Health Economics, 21,* 145-172. doi: 10.1002/hec.1697

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42. doi: 10.1257/089533005775196732

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axiom. *The Quarterly Journal of Economics, 75*(4), 643-669.

Hensher, D. A., Greene, W. H., Rose, J. M. (2005). *Applied choice analysis: A primer.* New York: Cambridge University Press.

Hole, A. R. (2008). Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. *Journal of Health Economics, 27,* 1078–1094.

Huber, J., Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research, 33,* 307-317.

Jadoul, M., Thumma, I., Fuller, D. S., Tentori, F., Li, Y., Morgenstern, H., Mendelssohn, D., Tomo, T., Ethier, J., Port, F., and Robinson, B. M. (2012). Modifiable practices associated with sudden death among hemodialysis patients in the dialysis outcomes and practice patterns study. Clinical Journal of the American Society of Nephrology, 7(5), 765-774. doi: 10.2215/CJN.08850811

Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B., & Bridges, J. F. P. (2013). Constructing experimental designs for discrete-choice experiments: Report of the ISPOR

conjoint analysis experimental design good research practices task force. *Value in Health, 16,* 3-13.

Kanheman, D., Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Klibanoff, P., Marinacci, M., Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, 73*(6), 1849-1892.

Loomes, G., Stramer, C., & Sugden, R. (1991). Observing violations of transitivity by experimental methods. *Econometrica, 59*(2), 425-439.

Loomes, G., & Sugden, R. (1982). Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal, 92,* 805-824.

Loomes, G., & Sugden, R. (1987). A more general form of regret theory. *Journal of Economic Theory 41,* 270-287.

Loomes, G., & Taylor, C. (1992). Non-transitive preferences over gains and losses. *The Economic Journal, 102,* 357-365.

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling, 3*(3), 57-72.
www.jocm.org.uk

Lancsar, E., Louviere, J. (2006). Deleting 'irrational' responses from discrete choice experiments: A case of investigating or imposing preferences? *Health Economics, 15,* 797-811.
doi:10.1002/hec.1104

Lloyd, A. J. (2003). Threats to the estimation of benefit: Are preference elicitation methods accurate? *Health Economics, 12* (5), 393–402
doi:10.1002/hec.772

Louviere, J. J., & Lancsar, E., (2009). Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law, 4,* 527-546.
doi:10.1017/S1744133109990193

Malinvaud, E. (1952). Note on Von Neumann-Morgenstern's strong independence axiom. *Econometrica, 20*(4), 679.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science, 5,* 275-279.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(15), 447-470.

NITp (2012). Note informative al paziente in attesa di trapianto di rene e rene-pancreas [Italian for "Information for patients of kidney and kidney-pancreas in waiting list"]. *NITp Website.* Retrieved from http://www.nitp.org/area_paziente/trapianti_carte_servizi/1202rene_rene-pancreas.pdf

Piattelli Palmarini, M. (2005). La cognizione dell'errore [Italian for "The cognition of the error"]. *Networks, 5, 128-142.*

Porteous, T., Ryan, M., Bond, & C. M. (2006). Preferences for self-care or professional advice for minor illness: a discrete choice experiment. *British Journal of General Practice, 56*(533), 911-917.

Ryan, M., & Bate, A. (2001). Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. *Applied Economics Letters, 8*(1), 59-63, doi: 10.1080/135048501750041312

Ryan, M., & Gerard, K. (2003). Using discrete choice experiments to value health care programmes: Current practice and future research reflections. *Applied Health Economics and Health Policy 2*(1), 55-64.

Ruggieri, G. (2014). Diffusione della dialisi peritoneale in Italia e nel mondo, suoi limiti e loro cause [Spread of peritoneal dialysis in Italy and in the world, limitations and causes]. *Giornale di Tecniche Nefrologiche & Dialitiche [Journal of Nephrological and Dyalitical Techniques], 26*(S5), S13-S14. doi: 10.5301/GTND.2014.11897

Savage, L. J. (1954). *The foundations of statistics.* New York: Dover Publications.

Samuleson, P. A. (1952). Probability, Utility, and the Independence Axiom. *Econometrica, 20*(4), 670-678.

San Miguel, F., Ryan, M., & Amaya-Amaya, M. (2005). 'Irrational' stated preferences: A quantitative and qualitative investigation. *Health Economics, 14,* 307–322. doi:10.1002/hec.912

Scott, A., Watson, M. S., & Ross, S. (2003). Eliciting preferences of the community for out of hours care provided by general practitioners: A stated preference discrete choice experiment. *Social Science & Medicine, 56*(4), 803-814.

Sen, A. (1977). Social choice theory: A re-examination. *Econometrica, 45*(1), 53-89.

Sen, A. (1993). Internal consistency of choice. *Econometrica, 61*(3), 495-521.

Stanovich, K. E., West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645-726.

Street, J. S., Burgess, L., Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing, 22,* 459–470.

Swait, J., & Adamowicz, W. (2001). Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes, 86,* 141-167.
doi:10.1006/obhd.2000.2941

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34,* 273-286.

Train, K. (2002). Discrete choice methods with simulation. Cambridge University Press.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior (2nd Edition).* Princeton: Princeton University Press.

Watson, V., Ryan, M., Brown, C. T., Barnett, G., Ellis, B. W., & Emberton, M. (2004). Eliciting preferences for drug treatment of lower urinary tract symptoms associated with benign prostatic hyperplasia. *The Journal of Urology, 172,* 2321-2325.
doi: 10.1097/01.ju.0000140957.31325.7f

Xuanming, S., & Zenios, S. A. (2006). Recipient choice can address the efficiency-equity trade-off in kidney transplantation: A mechanism design model. *Management Science, 52*(11), 1647-1660.
http://dx.doi.org/10.1287/ mnsc.1060.0541

Xuanming , S., Zenios, S. A., & Chertow, G. M. (2004). Incorporating recipient choice in kidney transplantation. *Journal of the American Society of Nephrology, 15*(6), 1656–1663.

Giacomo Battiston

30/10/2015

ERRATA CORRIGE

A further check on the data revealed that findings about fatigue effects were brought about by a small error in the analysis of data. In the end, no such effect emerges from the regressions effected on the data. Please, disregard the parts in the thesis dealing with them (they are highlighted in red in the text).

The sentence “Specifically, one of them had one item non-responded out of 16, while the other had 15” should read (p. 15) “Specifically, five of them had one item non-responded out of 16, while the other had 15” (The sentence is highlighted in yellow in the text.