Working papers series

WP ECON 08.010

Sequencing Anomalies in Choice Experiments

Brett Day (University of East Anglia).

Jose Luis Pinto Prades (U. Pablo de Olavide)

JEL Classification numbers: Q51, C14, I10

Keywords: Choice experiments; sequencing anomalies; ordering effects; dichotomous choice contingent valuation; non-parametric testing.

Department of Economics

Sequencing Anomalies in Choice Experiments

Brett Day[1,*] and Jose-Luis Pinto[2]

[1] Centre for Social and Economic Research on the Global Environment (CSERGE), School of Environmental Sciences, University of East Anglia, United Kingdom.

[2] Departamento de Economia, Métodos Cuantitativos e Historia Económica, Universidad Pablo de Olavide, Sevilla, Spain

* **Corresponding author**

Address: Brett Day, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom

e-mail: brett.day@uea.ac.uk

tel.: +44 (0)1603 591413

fax.: +44 (0)1603 507719

Submitted to JEEM: 10th September, 2008

Sequencing Anomalies in Choice Experiments

Abstract

This paper investigates whether responses to choice experiments (CEs) are subject to sequencing anomalies. While previous research has focussed on the possibility that such anomalies relate to position in the sequence of choice tasks, our research reveals that the particular sequence of tasks matters. Using a novel experimental design that allows us to test our hypotheses using robust nonparametric statistics, we observe sequencing anomalies in CE data similar to those recorded in the dichotomous choice contingent valuation literature. Those sequencing effects operate in both price and commodity dimensions and are observed to compound over a series of choice tasks. Our findings cast serious doubt on the current practice of asking each respondent to undertake several choice tasks in a CE whilst treating each response as an independent observation on that individual's preferences.

Keywords: Choice experiments; sequencing anomalies; ordering effects; dichotomous choice contingent valuation; non-parametric testing

JEL Classifications: Q51, C14, I10

1. Introduction

Over recent years, techniques of choice modelling such as choice experiments (CEs) have enjoyed a startling rise in popularity amongst the practitioners of non-market valuation (Adamowicz, 2004). The fundamental building block of a CE is a choice task. A choice task confronts subjects with two or more options where the options differ both in the qualities of the non-market good (the "commodity dimension") and in the cost imposed on the subject (the "price dimension"). The usual procedure is to ask subjects to indicate their preferred option in a series of such choice tasks. As such, CEs provide a rich data source from which researchers can deduce how subjects are prepared to trade off between money and the various dimensions of the commodity space.

In contrast, dichotomous-choice (DC) contingent valuation techniques, that had previously enjoyed the status of "most preferred valuation method", provide a relative paucity of data. In their earliest inception (Bishop and Heberlein, 1979), DC contingent valuation exercises presented each respondent with just one task; a choice between the status quo and the provision of the non-market good at a cost. Across a sample of respondents, the commodity dimensions were held constant whilst the price dimension was varied so that the willingness to pay distribution for that particular manifestation of the non-market good could be estimated. Whilst this 'single-bounded' DC (SBDC) elicitation method was strongly endorsed by the US National Oceanographic and Atmospheric Administration's blue ribbon panel (Arrow et al., 1993), practitioners were concerned by its relative inefficiency. In particular, the data from a SBDC cannot provide values for other manifestations of the non-market good. Moreover, since each respondent provides just one piece of information large samples sizes are required.

In response to the latter criticism, Hanemann, Loomis and Kanninen (1991) proposed the 'double-bounded' DC (DBDC) elicitation method. Here, following the initial DC question, a 'follow-up' DC question is asked offering the non-market good at a different price. The elicitation of a second response yields substantial gains in statistical efficiency.

Subsequent empirical testing, however, has revealed a robust anomaly in responses to DBDC questions. In particular, numerous studies have observed that the preferences implied by responses to first questions differ systematically from those implied by responses to follow-up questions (e.g. McFadden, 1994; Herriges and Shogren, 1996). Interestingly, the particular order of prices appears to be important; sequences of questions in which the price rises from the first question to the second have a much more profound impact on preferences than those that fall (De Shazo, 2002). The existence of these price-sequencing effects casts serious doubt on the validity of responses to follow-up questions. Indeed, one might argue that these well-documented problems have precipitated a growing disaffection with DC contingent valuation and contributed to the growing interest in CE methods.

While there are a number of differences between DC contingent valuation and CE approaches, there are also many similarities. For example, whilst differing in presentation, the SBDC elicitation method is essentially a simple form of CE in which subjects face only one choice task requiring a preference to be stated between the status quo and an alternative in which a non-market good is provided at a price. Likewise, the DBDC elicitation method is a CE with two choice tasks pitting the status quo against an alternative offering a non-market good. In this case, moving from the first choice task to the second, the price dimension of the alternative is altered, but there is no change in the commodity dimension.

Given these similarities a fundamental question that must be asked of the CE method is whether it is subject to the same price-sequencing anomalies as have been observed in DBDC contingent valuation studies. This paper sets out to address that question. In addition, since CEs allow commodity as well as price dimensions to vary across tasks, this study seeks to establish if equivalent commodity-sequencing anomalies are observable in CE responses. Indeed, our study is designed to examine how preferences are impacted in cases characterised by simultaneous price- and commodity-sequences. Finally, since CEs typically present a series of choice tasks, a further objective is to explore how sequencing anomalies develop over multiple tasks.

While a number of previous studies have investigated the issue of sequencing anomalies in CEs, most have sought to identify patterns of changes in stated preferences that relate simply to position in the sequence of tasks. In the main, this literature has been based on parametric modelling of preferences in the random utility framework. For example, Bateman et al. (2008) allow the coefficient of the preference function relating to the money attribute to vary according to task order and find that respondents' express a gradually higher marginal utility of income as they progress through the sequence of choice tasks. In a similar vein, a number of studies have parameterised the scale of the random element of the preference function so as to examine trends related to position in the question sequence (e.g. Adamowicz et al., 1998, Caussade et al., 2005). Evidence of increasingly variable preferences are usually interpreted as indicating respondent fatigue (e.g. Phillips, Johnson, and Maddala, 2002) while decreasing variability is usually interpreted as preference learning (e.g. Bateman et al., 2008, Holmes and Boyle, 2005).

In contrast, our work draws on the insights provided by the extensive literature on sequencing effects in DC contingent valuation. Rather than assuming that sequencing effects relate simply to position in the sequence, our work seeks to understand how sequencing effects are determined by the particular order in which the choice tasks are presented. As we shall see, we find that expressed preferences are profoundly affected by the particular sequence of tasks.

Within the context of CEs the study that is most similar to our work is that of Holmes and Boyle (2005). They test to see whether responses to CE tasks are influenced by the nature of the preceding and following choice tasks. They estimate a parametric model of preferences that includes terms reflecting the values of price and commodity attributes of options in those previous and future choice tasks. They find a number of these lagged terms to be significant determinants of preferences.

The study presented in this paper explores sequencing effects in a radically different testing framework. Our experiment employs a multiple split-sample design with each sample receiving a sequence of three choice tasks. The order of choice tasks differs across samples allowing us to examine a variety of hypotheses regarding the nature of sequencing effects. The design is such that each hypothesis can be tested through a simple across-sample comparison of responses to a particular choice task using robust non-parametric statistics. In contrast to previous studies, we are able to examine sequencing effects in CEs without imposing assumptions regarding the structure of preferences. Our nonparametric approach is both simple and transparent and as such, we believe, increases the credibility of our findings.

In the next section we describe the sorts of sequencing anomalies previously observed in contingent valuation exercises and consider how these might translate to CE-style questions. Section 3 describes the experiment we have designed to test those hypotheses. Section 4 presents the results of our empirical research and Section 5 provides a summary and some concluding remarks.

2. Sequencing Anomalies in Dichotomous Choice Contingent Valuation Studies

The standard neoclassical economic model asserts axiomatically that individuals have complete and coherent preferences. In the context of a CE that assertion dictates that respondents can always determine which of any set of options in a choice task is their most preferred. Moreover, provided they are motivated to answer truthfully then their responses should not be affected by features of the elicitation procedure that are, according to the theory, decision-irrelevant. For example, according to this model of behaviour the preferred option in a particular choice task should not change according to the nature of the options presented in previous choice tasks. Responses that conform to this prediction are said to demonstrate *procedural invariance*.

In the majority of applications of the CE methodology procedural invariance is assumed without testing. Violations of procedural invariance, however, present a profound problem for the elicitation of preferences using CEs. If, for example, expressed preferences differed systematically according to the sequence in which choice tasks were presented, then it would be impossible to assert that preferences recovered from a particular sequence of tasks were somehow ‘true’ preferences.

Of course, violations of procedural invariance are well-documented in the contingent valuation literature. In particular, several studies have reported *price-sequencing* anomalies in DBDC elicitation (e.g. McFadden, 1994; Cameron and Quiggen, 1994; Carson *et al*., 1994; Herriges and Shogren, 1996; Alberini, Kanninen and Carson, 1997; Bateman et al. 2001) possibly the most detailed of which being that of De Shazo (2002). De Shazo examined acceptance rates when a non-market good was offered at a particular price in the first question as compared to when it was offered at the same price in the follow-up question. In cases where those follow-up questions were preceded by a higher price (what we shall term an *improving price sequence*) there was no systematic difference in acceptance rates. However, when the preceding price was a lower amount (what we shall term a *worsening price sequence*) acceptance rates were significantly depressed.

If price-sequencing affects acceptance rates in DBDC contingent valuation exercises, then an obvious question would be to ask if *commodity-sequencing* also precipitates violations of procedural invariance. Bateman and Brouwer (2006) in their investigation of scope insensitivity in DC contingent valuation studies provide some insight into the possible nature of commodity-sequencing anomalies. They confront respondents with two SBDC questions concerning a low and high level of provision of a non-market good. They compare the median value estimated from a sample responding to the low provision question first with those estimated from a sample responding to the high provision question first. They observe what they describe as “some fanning out of estimates as we move from first to second responses” (p.207). That is to say, the implied values of the large and small levels of provision are relatively more similar when calculated from the first question responses than when calculated from the follow-up responses. This observation is consistent with a commodity-sequencing anomaly in which a good is regarded more favourably when preceded by a question offering a relatively smaller level of provision (an *improving commodity* *sequence*) whilst being regarded less favourably if preceded by a relatively larger level of provision (a *worsening commodity sequence*).

We summarise the evidence for sequencing effects from the DC contingent valuation literature in Table 1.

[INSERT TABLE 1 AROUND HERE]

Now, consider a simple CE in which the respondent is faced by a series of tasks each requiring a choice to be made between just two options (which we label Option I and Option II). To simplify further, imagine that the commodity dimension of the good can be described as either small, medium or large. Likewise the prices can take the values €0, €Low and €High. In this context, we can construct pairs of choice tasks that exactly replicate contingent valuation questions with price- and commodity-sequences.

For example, the sequence of choice tasks shown in the upper part of Figure 1 replicates an improving price sequence in a DBDC contingent valuation question. Here Option I is the same in both tasks (the status quo in a DBDC contingent valuation question). Likewise, the commodity dimension of Option II is the same in both tasks (the new level of provision of the non-market good in a DBDC contingent valuation question). All that changes across the choice tasks is that the price dimension of Option II falls (an improving price sequence in a DBDC contingent valuation question). Accordingly, we describe Option II in Choice Task 2 as appearing in an *improving price sequence*. Notice that in the context of a CE the sequence is attributed to the particular Option in which it is observed.

[INSERT FIGURE 1 AROUND HERE]

The lower part of Figure 1 shows a sequence of CE questions that replicates the worsening price sequence of a DBDC contingent valuation question. Again Option I is the same in both tasks. Option II also offers the same level of provision but at a low price in the first choice task followed by a high price in the second choice task. Again we attribute the sequence to the option such that we describe Option II in Choice Task 2 as appearing in a *worsening price sequence*.

If the series of choice tasks in Figure 1 exactly replicate improving and worsening price sequences in DBDC contingent valuation questions, then we might expect them to generate the same sequencing anomalies observed by De Shazo (2002). We could simply test that contention by presenting independent samples with the two sequences of choice tasks. Observe that the first choice task faced by one sample is identical to the second choice task faced by the other sample. Procedural invariance demands that the two samples will provide roughly similar responses to this task. In contrast, De Shazo's results suggest that procedural invariance will only characterise the improving price sequence. The worsening price sequence, it is predicted, will induce a sequencing anomaly. In particular, the proportion favouring Option II when presented in the second choice task in a worsening price sequence will be significantly less than the proportion favouring that Option when presented in the first task to the other sample.

Figure 2 presents a similar construction but this time illustrating Option I of the choice tasks following improving and worsening commodity sequences. In the upper diagram, for example, Option II does not change and in Option I the price is the same but the amount of good increases. This is then an *improving commodity sequence* in Option I. The opposite pattern, a *worsening commodity sequence* in Option I, is illustrated in the lower part of Figure 2.

[INSERT FIGURE 2 AROUND HERE]

Again, the first choice task in the upper sequence is identical to the second choice task in the lower sequence. As such, a simple test for procedural invariance would be to present the two sequences of choice tasks to independent samples and compare response proportions for the matched tasks. If the observations of Bateman and Brouwer (2006) carry over to the CE framework, then we expect procedural invariance to be violated in both improving and worsening commodity sequences in the directions indicated in Table 1.

Of course, in a CE it is unusual to have only one of the options changing from choice task to choice task. For example, Figure 3 illustrates choice tasks that present respondents with options in mixed commodity and price sequences; that is to say, one option presents an improving (worsening) price sequence whilst the other an improving (worsening) commodity sequence. The possibility exists that sequencing anomalies might arise in this context also.

[INSERT FIGURE 3 AROUND HERE]

If the patterns of behaviour observed in DC contingent valuation studies carry over to these more complex choice situations, then the improving mixed sequence should result in violations of procedural invariance. In particular, in the improving sequence the relatively larger commodity offered by Option I in the second task leads respondents to regard this option more favourably, whilst the improved price offered by Option II in the second task results in no such equivalent bias. This combination of effects would lead to a relatively greater proportion of respondents favouring Option I in this choice task than would do if that same choice task was the first in the sequence.

In the worsening mixed sequence, the two sequencing biases work in the same direction; the relatively smaller commodity offered by Option 1 makes this option appear less favourable but the relatively greater price offered by Option 2 makes this also appear less favourable. Since we are unable to determine in advance which of the two sequencing effects will dominate, it is not possible to make predictions concerning violations of procedural invariance in this case.

3. Experimental Design and Testing Framework

Our application concerns the valuation of health using a CE. In each choice task respondents were asked to imagine that they had been diagnosed with a medical problem that would result in a considerable deterioration in their quality of life. Quality of life was measured using the Euroqol (EQ-5D) (Brooks, 1996). EQ-5D is a standardised instrument for use as a measure of health outcome. It describes any health state through five dimensions; mobility, self-care, usual activities, pain/discomfort, anxiety/depression. Each dimension can take one of three levels; no problems, some or moderate problems, extreme problems. For example, health state 11111 implies full health since it is equivalent to having no problems in each of the five dimensions. However, health state 33333 implies having extreme problems in each of the five dimensions. The advantage of using EQ-5D is that there are clear dominance relations between health states. In this study we used two EQ-5D health states, namely, 21212 and 22223. Cleary, an improvement from 21212 to 11111 (full health) implied a smaller health gain than an improvement from 22223 to full health. A summary of the health state descriptions is provided in Figure 4.

[INSERT FIGURE 4 AROUND HERE]

Respondents were informed that the problem was treatable and that following treatment they would be returned to full health within one year. In each choice task, subjects were asked to choose between two treatment options. One option was a medicine provided by the hospital. While this option was free of charge it meant that the subject would still experience quality of life-reducing symptoms for the first 2 months of treatment. The second option in each choice task was to purchase an alternative medicine from a pharmacy. While this treatment was costly, it meant that the subject would avoid any symptoms whilst being treated thereby enjoying a quality of life that was equivalent to full health.

The commodity dimension of options in our CE described one of three states of health; two months of severe ill-health (health state 22223) or two months of mild ill-health (health state 21212) or on-going full health (health state 11111). In addition, the price dimension of options in our CE described one of four levels of treatment cost; €0, €60 (€Low), €120 (€Med) or €240 (€High). As shown in Figure 5, a typical choice task pitted a zero cost treatment with 2 months of reduced quality of life, against a costly treatment that returned the subject to immediate full-health.

[INSERT FIGURE 5 AROUND HERE]

In order to simplify the presentation of our experimental design, it is expedient to further condense each choice task into the simple schematic shown on the right hand side of Figure 5. Here the top box represents Option I where the cost, treatment duration and symptom's duration dimensions have all been suppressed since these were always €0, 12 months and 2 months respectively. Likewise the bottom box represents Option II with the commodity dimensions suppressed since these were always 12 months of treatment with no symptoms.

For the purposes of our study, we constructed four different choice tasks. Our experimental design involved a six-way split sample with each sample facing a different set of three of those four choice tasks. We label the six samples, A_1, A_2, B_1, B_2, C_1 and C_2. The tasks and order in which they were presented to respondents in each sample are summarized using the simplifying schematic in Figure 6.

[INSERT FIGURE 6 AROUND HERE]

Procedural invariance requires that two independent samples presented with the same choice task will express approximately the same preferences irrespective of the details of any preceding choice tasks. The essence of our experimental design, therefore, is to confront independent samples with the same choice task but set in differing sequences and to record the proportions choosing each option. In this context, the null hypothesis can be tested using robust nonparametric tests to compare those proportions across the two samples.

The first two tasks of our design are selected to test whether elicitation of preferences using CE-style questions precipitates the same sequencing anomalies observed in DC contingent valuation. For clarity, Table 2 describes the sequences for both options in these first two tasks and indicates the bias those sequences would precipitate should the sequencing effects observed in DC contingent valuation studies carry over to the CE context. Observe that the commodity on offer in our experiment is a bad. As such, an improving commodity sequence is one that moves from a worse health state (severe) to a better health state (mild).

The comparison of response proportions for the first two choice tasks of samples A_1 and A_2 provide tests of price-sequencing (as per Figure 1). The same comparison for samples B_1 and B_2 test for commodity-sequencing (as per Figure 2). While the comparison of samples C_1 and C_2 test for mixed price- and commodity-sequencing (as per Figure 3).

[INSERT TABLE 2 AROUND HERE]

Through the inclusion of a third task, our design also allows us to consider how sequencing anomalies might evolve over a series of tasks. We return to discuss this issue in more detail in the next section.

4. Results

Each of the six samples was chosen so as to provide a representative sample of the population of Northern Spain that formed the location of interest for our investigations. Surveying was undertaken by professional interviewers in personal interview sessions. Each sample contained 83 observations except for sample B_2 which contained 85. The data are summarised in Figure 6 which shows the proportions of each sample choosing each option in each choice task. For ease of exposition, we refer to the first question faced by sample A_1 as A_{11}, likewise the second question faced by this sample is A_{12}, and so on.

4.1 Tests of Consistency

The fundamental contention of our experimental design is that each sample represents an independent observation of the same underlying population of preferences. Only if that is the case will it be sensible to draw inference from across-sample comparisons of responses. Moreover, we would hope that the preferences expressed by individuals in those samples conform to some basic tenets of economic theory. If this were not true, then we might conclude that responses to our choice tasks provide no meaningful information on economic behaviour. Our first set of tests are designed to confirm (or refute) these fundamental assertions.

a. Tests 1 to 3: Across-sample consistency in identical first choice task

Each sample received identical information up to the point of the first choice task. Uniquely, therefore, responses to first choice tasks are unaffected by potential sequencing biases. We use that fact to test the contention that each sample is an independent observation of the same population preferences.

Samples A_1, B_2 and C_2 were presented with identical first choice tasks (tasks A_{11}, B_{21}, and C_{21} respectively). Our expectation is that the distribution of preferences in the three samples should be approximately the same. What we observe is that 25.3% of sample A_1, 25.9% of sample B_2 and 28.9% of sample C_2 chose Option I in the first task. As shown in Tests 1 to 3 of Table 3, a series of pairwise comparisons using a two-tailed chi-squared test of differences in proportions confirms that these differences are not statistically significant.

Encouragingly, for these three samples at least, our choice tasks appear to be tapping into a distribution of preferences that does not differ systematically across samples. For the purposes of subsequent testing, we combine these three observations and treat them as observations pertaining to the same set of underlying preferences. Of that combined sample of 187 individuals, 25.5% chose Option I.

[INSERT TABLE 3 AROUND HERE]

b. Test 4: Across-sample consistency - more commodity at same price

Standard economic theory indicates that individuals prefer more of good things to less. Accordingly, all else equal, we would expect more individuals to choose a treatment if it offered a greater health benefit.

Samples A_1, B_2, C_2 and sample B_1 face a first choice task in which the treatment promising an immediate return to full health has the medium price. For samples A_1, B_2 and C_2 the alternative treatment entails suffering an episode of severe ill-health while for sample B_1 the alternative treatment involves an episode of only mild ill-health. In line with expectations, the proportion choosing this alternative jumps from 25.5% when the ill-health event is severe (A_{11}, B_{21} and C_{21}) up to 37.3% when the ill-health event is mild (B_{11}). As shown in Test 4 of Table 3, a one-tailed test of differences in proportions confirms this to be a statistically significant difference (p-value: 0.0190).

c. Tests 5 & 6: Across-sample consistency - same commodity at different price

In a similar vein, economic theory postulates that when making purchases people prefer to pay less rather than more. All else equal, we would expect fewer individuals to choose a treatment the higher its price. Again, that prediction is testable by comparing response proportions across samples.

Samples A_1, B_2, C_2 and sample A_2 face a first choice task in which the costless treatment involves enduring a severe ill-health event. For samples A_1, B_2 and C_2, the alternative treatment, promising an immediate return to full health, has the medium price while for sample A_2 this treatment is offered at the high price. In line with expectations, the proportion choosing this alternative falls from 74.5% when offered at the medium price (A_{11}, B_{21} and C_{21}) to 67.5% when offered at the high price (A_{21}). A similar comparison is possible between sample C_1 and B_1 both of whom face a first choice task in which the costless treatment results in a mild ill-health event. For sample C_1 the alternative treatment can be obtained at the low price while for sample B_1 that treatment is offered at the medium price. Again, we observe what we expect; 77.1% choose this alternative at the low price (C_{11}) compared to 62.7% at the high price (B_{11}). Referring to Tests 5 and 6 in Table 3, a one-tailed test reveals the first of these comparisons to be insignificant (p-value: 0.1061), while the second is highly significant (p-value: 0.0212).

d. Within-subject consistency

One further expectation of the standard economic model is that individuals respond to repeated choice tasks with reference to a stable set of well-formed preferences. Under that assumption, we would expect respondents' choices to exhibit internal consistency. For example, it would not be consistent for an individual in sample C_1 to refuse to pay the low price for immediate full health in the first choice task (C_{11}) but then agree to pay the medium price in the third choice task (C_{13}) when in both tasks the alternative costless treatment results in the same mild ill-health event.

Respondents to the survey, it transpires, are remarkably consistent. In all 151 cases where individuals made choices in the first task that reveal preferences dictating a particular choice in the second task, that consistent choice was made. Of the 138 cases where a choice in the first task dictates a particular choice in the third task, only 4 individuals make the inconsistent choice.

So far, the prognosis for CE is rather positive. Respondents achieve a high level of internal consistency in their responses and across-sample tests suggest those responses conform to some basic tenets of economic theory. Importantly, a test of responses to the first choice task finds the distribution of preferences expressed by different samples to be statistically indistinguishable. Accordingly, we take those first task responses as unbiased reflections of population preferences and treat responses to the first choice tasks as *control cases* to which other responses can be compared.

4.2 Price Sequence Effects

The stylised facts from the contingent valuation literature suggest an asymmetric response to price sequences (see Table 1). If the CE elicitation format induces the same systematic shifts in expressed preferences then we would expect to see unusually low acceptance rates for an option in a worsening price sequence but no significant shifts for options in improving price sequences. As shown in Table 2, a comparison of the responses of samples A_1 and A_2 provide a test of that contention.

a. *Test 7: Worsening price sequence*

The control case for our worsening price sequence test is provided by sample A_2. Here we observe 67.5% of the sample choosing Option II in the first task (A_{21}). Sample A_1 face the exact same task but as their second choice (A_{12}). This second task differs from the first (A_{11}) only in so much as the price of Option II increases. This worsening price sequence has a substantial impact with only 57.8% of that sample selecting Option II, almost 10% less than in the control case. As shown in Test 7 of Table 4, a one-tailed test reveals this difference to be significant at the 90% level of confidence (p-value: 0.0996).

[INSERT TABLE 4 AROUND HERE]

b. *Test 8: Improving price sequence*

The control case for our improving price sequence test is provided by samples A_1, B_2 and C_2 that share the same first choice task. In response to those tasks (A_{11}, B_{21} and C_{21}), we observe 74.5% of respondents choosing Option II. The exact same choice is presented to sample A_2 as their second task (A_{22}). This second task differs from the first (A_{21}) only in that Option II is offered at a lower price. This improving price sequence has a negligible impact on expressed preferences with the proportion selecting Option II remaining relatively stable at 75.9%. Not surprisingly, Test 8 of Table 4 shows that these two proportions do not differ significantly (p-value from a two-tailed test: 0.7896).

Our data provide evidence of price-sequencing effects in CE tasks. We observe a substantial (if only marginally significant) reduction in the frequency with which respondents choose an option if, all else equal, that option is in a worsening price sequence. In contrast, no systematic impact on preferences can be discerned for options in improving price sequences. The pattern we observe in our CE data exactly replicates the sequencing effects recorded by De Shazo (2002) in DBDC contingent valuation data.

4.3. Commodity Sequence Effects

The stylised facts from the contingent valuation literature (Table 1) point to the possibility that an option in a worsening (improving) commodity sequence will be regarded less (more) favourably by respondents. As shown in Table 2, the responses of samples B_1 and B_2 provide the basis for our tests of those possible commodity-sequencing effects.

a. Test 9: Worsening commodity sequence

The control case for our test of worsening commodity sequences is provided by the first choice task presented to samples A_1, B_2 and C_2. Some 25.5% of that combined sample chose the severe ill-health event offered by Option I. Sample B_1 face the exact same task (B_{12}) but second in the sequence where the previous task (B_{11}) differed only in that Option I entailed enduring a mild ill-health event. This worsening commodity sequence has a major impact on the selection of Option I with only 8.4% of the sample making that choice. As shown in Test 9 of Table 4, a one-tailed test reveals this to be a highly significant difference (p-value: 0.0002).

b. Test 10: Improving commodity sequence

The control case for the improving commodity sequence is provided by the first task presented to sample B_1. We observe 37.7% of that sample plumping for the mild ill-health event offered by Option I. Sample B_2 faced the same task second in the sequence such that Option I is in an improving commodity sequence. Again this precipitates a considerable shift in expressed preferences with the proportion choosing Option I jumping to 64.7%.The formal test shown as Test 9 in Table 4 confirms this to be a highly significance difference (*p*-value from a one-tailed test: 0.0005).

Accordingly, the pattern of responses observed in our data offers strong evidence of commodity-sequencing anomalies. Moreover, in contrast to the asymmetry observed in price-sequencing, these anomalies function in both improving and worsening directions. Again, our data confirm that the patterns of commodity-sequencing anomalies observed in DC contingent valuation studies persist in CEs.

4.4. Mixed Sequence Effects

Of course, unlike DC contingent valuation, CEs typically vary more than just one attribute of one option from choice task to choice task. Such variation may lead to situations in which both options in a task are in sequences that might potentially influence respondents' choices. We refer to such tasks as being in a mixed sequence. As described in Table 2, samples C_1 and C_2 provide the basis for our tests of the biases induced by mixed sequences.

a. *Test 11: Improving mixed price and commodity sequence*

Consider first the comparison between responses to questions C_{11} and C_{22}. Both present a choice between a costless treatment resulting in a mild ill-health event (Option I) and a low priced treatment that removes all symptoms (Option II). While C_{11} represents our control case, C_{22} is the second choice task such that Option I is in an improving commodity sequence (the previous task offering a severe ill-health event) and Option II is in an improving price sequence (the previous task offering a medium price). Given our earlier findings, our expectation is that respondents to task C_{22} will be much better disposed to Option I because it is in an improving commodity sequence but that the improving price sequence will have no impact on preferences for Option II. Accordingly we might expect rather more respondents to select Option I in C_{22} than in the control case (C_{11}). As shown in Test 11 of Table 4, this is what we observe; 22.9% prefer Option I in C_{11} but this proportion increases to 41% in C_{22}. A one-tailed test reveals this to be a significant difference (*p*-value: 0.0063).

b. Test 12: Worsening mixed price and commodity sequence

The worsening mixed price and commodity sequences are tested by comparing responses to C_{12} to the control case provided by A_{21}. Here we have no directional hypothesis since in task C_{12} both options are in worsening sequences and hence both should be regarded less favourably. The data provides support for this contention as in both tasks the proportion favouring Option I is identically 32.5%. As shown in Test 12 of Table 4, there are no statistically significant differences in these observations.

It appears that the price and commodity-sequencing effects we have documented in very simple sequences (where the attributes of just one option of the choice change) can be used to explain the patterns of response observed in these more complex mixed sequences (where the attributes of both options of the choice change). In particular, the sequencing effect in a mixed sequence can be explained as the compounded impacts of the commodity and price-sequencing effects acting on the individual options of the choices.

4.5. Third Task Sequence Effects

Thus far our examination has focused on the nature of the sequence of first and second choice tasks. We shall refer to that sequence as the *initial sequence*. Moreover, since the response anomalies observed in the second task are determined by the nature of the immediately preceding task we shall dub those observed anomalies *first-order sequencing effects*. In Figure 7 the arrow indicating the first-order initial sequence is depicted as a solid line to reflect that this sequence precipitates *observed* response anomalies in the second choice task.

When a third choice task is introduced, further sequences are identifiable. In particular, the third task forms a first-order sequence with the second task; a sequence that we shall refer to as the *secondary first-order sequence*. In addition, the third choice task forms *a second-order sequence* with the first choice task. In Figure 7, these sequences are depicted as dashed arrows to reflect the fact that we have yet to establish whether they precipitate response anomalies in the third choice task.

[INSERT FIGURE 7 AROUND HERE]

One possibility is that respondents react to both or either of these additional sequences using the same heuristics that drove the biases we have observed in the initial first-order sequence (replicated for reference in the second column of Table 5). Under that assumption, we can identify the particular type of sequence characterising the move from the second to third choice tasks and thereby predict the direction of bias the secondary first-order sequence might precipitate in responses to the third task. Those predictions are listed in column 3 of Table 5.[1] Likewise, comparing the first task with the third task allows us to predict the direction of bias that might be precipitated by the second-order sequence. Those predicted biases are listed in column 4 of Table 5.

To illustrate, the initial first-order sequence for sample A_1 is a worsening price sequence that we have observed to shift preferences in favour of Option I. That observed bias is indicated by a plus sign in column 2. The secondary first-order sequence for sample A_1 is an improving mixed sequence. As an initial sequence that pattern was observed to bias

[1] In making these predictions, recall that the sequencing effect precipitated by the worsening mixed sequence does not provide a unique directional prediction. Rather the overall effect depends on the relative strength of conflicting improving price and commodity sequence effects. The initial sequence of Sample C_1 is a worsening mixed sequence. In this sequence it appears that the impact of the worsening price sequence in Option II, moving from Low to High, cancels out the impact of the worsening commodity sequence in Option I, moving from Mild to Severe. In that case, no response bias is seen in the second task. The secondary sequence of Sample B_2 is a similar worsening mixed sequence. Option I is in the identical commodity sequence (mild to severe) but the worsening price sequence in Option II is less extreme (Med to High). In this case, it seems unlikely that the two sequencing effects will cancel. Accordingly, the entry in column 3 of Table 5 for sample B_2 contains a minus symbol.

preferences in favour of Option I (sample C_2). Accordingly, we record that prediction in column 3 with a plus symbol. Finally, the second-order sequence for this sample forms an improving commodity sequence. As an initial sequence that pattern of tasks biases responses in favour of Option I (sample B_2). As such, column 4 for sample A_2 contains a plus symbol.

Columns 5 to 8 of Table 5 provide a summary of the direction of biases actually observed in the third choice tasks. In each case, the response proportions of the third task have been compared to the appropriate control case and one- and two-tailed tests of differences in proportions performed in order to identify the statistical significance of any differences. For example, in the case of Sample A_1 we observe responses to the third task biased in favour of Option I. Of course, that bias is what we might expect given that all three sequences for that sample are predicted to push preferences in that direction. Indeed, the bias is shown to be significant at the 90% level of confidence using a one-tailed test.

Unfortunately, since for sample A_1 all sequences work to bias responses in the same direction, that sample is relatively uninformative in untangling which of the sequences, if any, are responsible for the response bias observed in the third task. The possibility remains that responses to the third task are influenced by any combination of the various first- and second-order effects. In the following, we use our data to explore a number of hypotheses regarding the determinants of bias in third tasks.

[INSERT TABLE 5 AROUND HERE]

a. *Hypothesis 1: No Sequencing Effects in Responses to Third Choice Tasks*

We begin by establishing that response biases can be observed in the third choice tasks. After all, it might be the case that the sequencing effects seen in response to the second choice task result from respondents reacting adversely to the unexpected presentation of this subsequent question (‘a moment ago, you offered me that same commodity at a better price!’). It is possible, that by the time the third question is presented, respondents have come to terms with the repeated nature of the exercise, such that responses to the third task are not distorted by sequencing effects.

In this case, our alternative hypothesis is simply that there is some unspecified difference in responses between first and third choice tasks. As such, the appropriate test is two-tailed. Looking down the final column of Table 5, it is clear that in the cases of samples A_2 and C_2 we can reject the null of consistency of responses to first and third choice tasks with greater than 90% and 99% confidence respectively. Evidently third choice tasks are not immune from response anomalies.

b. Hypothesis 2: Residual Effect from Initial Sequence

Maintaining the assumption that secondary first-order and second-order sequences do not generate response anomalies, an alternative hypothesis asserts that the initial sequence can cause a systematic shift in respondents' preferences that persists in their consideration of the subsequent choice tasks. In this case, we would expect the observed response bias from the initial sequence (column 2 of Table 5) to have the same sign as the observed response bias in the third task (column 5 of Table 5). Since these present directional hypotheses a one-tailed test is appropriate.

Samples A_1 and C_2 provide responses that conform to this hypothesis. The initial sequence precipitates a positive bias in favour of Option I and we also observe significant bias in the same direction in the third task. The responses of the remaining samples, however, provide contradicting evidence. The preferences of samples A_2 and C_1 are apparently unaffected by the initial sequences yet we observe statistically significant biases in favour of Option I in responses to the third task. Likewise, samples B_1 and B_2 express preferences that are significantly affected by the initial sequence but responses to the third task appear unbiased. The hypothesis of responses being influenced solely by the initial sequence is firmly rejected by our data.

c. Hypothesis 3: Second-Order Effect

Another possibility is that respondents anchor on the attribute levels in the first choice task. Subsequent tasks are compared with this initial task and it is this comparison which precipitates response anomalies. As we have seen, anomalies in responses to the second task can be explained through such a comparison. For third choice tasks the relevant comparison is that provided by the second-order sequence. As such, the response anomalies we might expect to observe in the third choice task are given by the directional predictions listed in column 4 of Table 5.

The data conform much more closely to the predictions of this behavioural theory. We find that for samples A_1, A_2, C_1 and C_2 the bias in the third choice task is in the direction of that predicted from the second-order sequences and in all cases a one-tailed test confirms these biases to be significant with 90% confidence or greater. Sample B_1 also conforms with the prediction though, in this case, that prediction is for no response anomaly in the third task. A two-tailed test confirms that there is no significant bias in responses to the third task. In the case of sample B_2, however, the second-order sequence would imply a bias in favour of Option I in third task responses. A one-tailed test rejects that hypothesis.

Accordingly, the proposition that the direction of response anomalies in third choice tasks can be explained through second-order sequence effects alone fails to completely explain the response patterns observed in our data.

d. Hypothesis 4: Compounding First-order Effects

A further possibility is that response anomalies are precipitated primarily by first-order sequencing effects. Over a series of choice tasks, responses may be influenced by each first-order sequence encountered. The initial sequence may systematically shift respondents' preferences and then from that baseline the secondary sequence may precipitate a further shift. According to this theory, the bias we would expect to observe in responses to third tasks should be in the direction suggested by a compounding of the two first-order sequence effects.

With reference to Table 5, we find that for all samples the predictions made by this hypothesis entirely conform to the observed direction of bias in third task responses. In sample A_1 the initial and secondary first-order sequences both bias responses in favour of Option I and a one-tailed test confirms that responses to the third task are significantly biased in the same direction (p-value: 0.0587). In samples A_2 and C_1, the initial sequence does not bias responses but the secondary sequence leads to a favouring of Option I. In both cases a significant bias in the predicted direction is observed in the third task responses (p-values: 0.0306 and 0.0587 respectively). In sample C_2 the initial sequence biases responses in favour of Option I but the secondary sequence should have no effect. Again we observe a significant bias in the expected direction (p-value: 0.0003). Finally, in samples B_1 and B_2, the initial and secondary sequences imply biases that operate to move preferences first in one direction and then in the other. As a result, the observation of no significant biases in third task responses (p-values: 0.8547 and 0.3947 respectively) entirely conforms to this behavioural model.

In addition to the direction of bias, the idea of compounding first-order effects has implications for the magnitude of observed sequencing anomalies. For example, the initial sequence faced by sample A_2 has no systematic impact on preferences. As such, we might expect the secondary sequence to induce the same magnitude of response anomaly as would be observed if it were the initial sequence. Observe that the secondary sequence faced by sample A_2 is identical to the initial sequence faced by sample C_2. If our contention is correct then the response proportions for task A_{23} should closely resemble those for task C_{22}. Again, the data supports the behavioural model; we observe 36.1% choosing Option I in task A_{23}, an amount that does not differ significantly from the 41% choosing that option in task C_{22} (p-value from a two-tailed test: 0.5326).

A contradicting example is provided by a comparison of samples A_1 and C_1. Here we would expect the magnitude of the response bias in the third choice task to be larger in A_{13} than in C_{13}. The reason for that expectation is that both samples face an identical secondary sequence. The initial sequence of sample A_1, however, acts to positively reinforce the bias of the secondary sequence whereas the initial sequence of sample C_1 is observed not to bias responses. In actuality, we observe the response bias in the third task to be of identical magnitude.

The theory of compounding first-order sequencing effects successfully organises our data. The predictions of that behavioural model perfectly predict the direction of biases observed in responses to the third choice task. Having said that, we acknowledge that we do not have a perfect explanation for all of our data; in one case we observe that the magnitude of response biases does not conform to expectations. In addition, our data does not allow us to identify whether the compounding of first-order sequencing effects is further compounded by the second-order sequence. A new experimental design would be needed to cast greater light on these issues.

5. Summary and Concluding Remarks

Despite the recent and rapid uptake of CE methods by non-market valuation practitioners, there are reasons to suspect that the method may suffer from sequencing anomalies. Previous research in this area has, in the main, focussed on the possibility that responses to CEs are affected by position in the sequence of tasks. In contrast, our research explores the question of whether the particular sequence of choice tasks matters. Drawing on the insights provided by the literature on sequencing anomalies in iterative DC contingent valuation studies this paper reports on an experiment specifically designed to assess that question.

Our findings categorically reject the assumption that CEs are immune from sequencing anomalies. Moreover, they confirm that the pattern of sequencing anomalies is determined by the particular sequence of choice tasks. What is more, those patterns closely replicate the patterns of sequencing anomaly observed in the DC contingent valuation responses. To be specific, our results indicate that the frequency with which respondents choose options in worsening price sequences is significantly reduced though the opposite anomaly is not observed for options in improving price sequences (a pattern observed in DC contingent valuation data by De Shazo, 2002). In addition, we find highly significant anomalies resulting from commodity sequences. In this case, the effects work in both directions with significant reductions (increases) in choice probabilities observed for options in worsening (improving) commodity sequences. An open question raised by our research is why significant sequencing effects are observed in improving commodity sequences but not improving price sequences.

The patterns of choice suggested by these sequencing behaviours completely organise our data, explaining the direction of bias observed to choice tasks in mixed sequences; that is, when one option of a choice task is presented in an improving (worsening) price sequence, while the other is presented in an improving (worsening) commodity sequence. In addition, we find robust evidence to suggest that these first-order sequencing anomalies are compounded over a series of choice tasks.

The central message of our paper is that CEs are vulnerable to sequencing anomalies and that the particular pattern of anomaly can be explained by the particular sequence of choice tasks. Our findings cast serious doubt on the current practice of asking each respondent to undertake several choice tasks in a CE whilst treating each response as an independent observation on that individual's preferences.

References

Adamowicz, W. L., (2004). "What's it worth? An examination of historical trends and future directions in environmental valuation", *Australian Journal of Agricultural and Resource Economics*, 48(3), pp. 419–443.

Adamowicz, W. L., P. Boxall, M. Williams and J. J. Louviere (1998). "Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation," *American Journal of Agricultural Economics*, 80, pp. 64-75.

Alberini, A., B. Kanninen and R. T. Carson (1997). "Modelling response incentive effects in dichotomous choice contingent valuation data", *Land Economics*, 73, pp. 309-324.

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner and H. Schuman (1993). "Report of the NOAA panel on contingent valuation", *Federal Register*, 58, pp.4601-4614.

Bateman, I. J., I. H. Langford, A. P. Jones and G.N. Kerr, (2001). "Bound and path effects in double and triple bounded dichotomous choice contingent valuation", *Resource and Energy Economics*, 23(3), pp. 191-213.

Bateman, I. J. and R. Brouwer (2006). "Consistency and construction in stated WTP for health risk reductions: A novel scope-sensitivity test", *Resource and Energy Economics*, 28, pp. 199-214.

Bateman, I. J., R. T. Carson, B. H. Day, D. Dupont, J . J. Louviere, S. Morimoto, R, Scarpa and P. Wang (2008). "Choice set awareness and ordering effects in discrete choice experiments", *CSERGE Working Paper*, EDM-2008-01, University of East Anglia.

Bishop, R. C., and T. A. Heberlein. (1979). "Measuring values of extramarket goods: Are indirect measures biased?", *American Journal of Agricultural Economics*, 61(5), pp. 926-930.

Brooks, R., (1996). "EuroQol: The current state of play", *Health Policy*, 37(1), pp. 53-72.

Cameron, T. A., and J. Quiggin (1994). “Estimation using contingent valuation data from a dichotomous choice with follow-up questionnaire”, *Journal of Environmental Economics and Management*, 27(3), pp. 218-234.

Caussade, S., J. de D. Ortuzar, L. Rizzi, and D. Hensher (2005). “Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates. *Transportation Research B*, 39, pp. 621–640.

De Shazo, J.R., (2002). “Designing transactions without framing effects in iterative question formats”, *Journal of Environmental Economics and Management*, 43, pp. 360-385.

Hanemann, W. M., J. Loomis, and B. Kanninen (1991). “Statistical efficiency of double-bounded dichotomous choice contingent valuation”, *American Journal of Agricultural Economics*, 73(4), pp. 1255-1263.

Holmes, T. P., and K. J. Boyle (2005). “Dynamic learning and context-dependence in sequential, attribute-based, stated preference valuation questions”, *Land Economics*, 81(1), pp. 114-126.

Herriges, J., and J. Shogren (1996). “Starting point bias in dichotomous choice valuation with follow-up questioning”, *Journal of Environmental Economics and Management*, 30, pp.112-131.

McFadden, D., (1994). “Contingent valuation and social choice”, *American Journal of Agricultural Economics*, 76, pp. 689-708.

Phillips, K. A., F. R. Johnson, and T. Maddala (2002). “Measuring What People Value: A Comparison of ‘Attitude’ and ‘Preference’ Surveys,” *Health Services Research*, 37, pp. 1659-1679.

TABLES

Table 1: Impact of sequence on acceptance rates as observed in dichotomous choice contingent valuation studies

		Sequence	
		Worsening	**Improving**
Dimension	**Price**	–	None
	Commodity	–	+

Table 2: Possible patterns of sequencing and predictions of resultant response bias in second choice task

Sample	Option	1st choice	2nd choice	Sequence	Prediction of bias in Second Choice Task
A_1	I	Severe	Severe	No sequence	Increase attractiveness of Option I
	II	€Med	€High	Worsening	
A_2	I	Severe	Severe	No sequence	No bias
	II	€High	€Med	Improving	
B_1	I	Mild	Severe	Worsening	Reduce attractiveness of Option I
	II	€Med	€Med	No sequence	
B_2	I	Severe	Mild	Improving	Increase attractiveness of Option I
	II	€Med	€Med	No sequence	
C_1	I	Mild	Severe	Worsening	No prediction
	II	€Low	€High	Worsening	
C_2	I	Severe	Mild	Improving	Increase attractiveness of Option I
	II	€Med	€Low	Improving	

Table 3: Tests of across-sample consistency in responses to Choice Task 1

Test	**Test Case**	**Control case**	**Expected Difference in Response Proportions**	**Signed *p*-value**
Response proportions the same for identical choice tasks:				
Test 1:	A_{11}	B_{21}	None	+0.9312[b]
Test 2:	A_{11}	C_{21}	None	+0.6004[b]
Test 3:	B_{21}	C_{21}	None	+0.6593[b]
More is better:				
Test 4:	A_{11}, B_{21} & C_{21}	B_{11}	+ Option I	+0.0190[a]
Higher prices are worse:				
Test 5:	A_{11}, B_{21} & C_{21}	A_{21}	– Option II	-0.1061[a]
Test 6:	C_{11}	B_{11}	– Option II	-0.0212[a]

Notes:

[a] From one-tailed test of equality of proportions

[b] From two-tailed test of equality of proportions

Table 4: Tests of Sequencing Anomalies

Sequence	Test Case	Control case	Expected Difference in Response Proportions	Signed *p*-value
Price Sequence:				
Test 7: Worsening	A_{12}	A_{21}	– Option II	-0.0996[a]
Test 8: Improving	A_{22}	A_{11}, B_{21} & C_{21}	None	+0.7896[b]
Commodity Sequence:				
Test 9: Worsening	B_{12}	B_{21}	– Option I	-0.0002[a]
Test 10: Improving	B_{22}	B_{11}	+ Option I	+0.0005[a]
Mixed Sequence:				
Test 11: Improving (price) Improving (commodity)	C_{22}	C_{11}	+ Option I None Option II	+0.0063[a]
Test 12: Worsening (price) Worsening (commodity)	C_{12}	A_{21}	- Option I - Option II	1.000[b]

Notes:
[a] From one-tailed test of equality of proportions
[b] From two-tailed test of equality of proportions

Table 5: Direction of first- and second-order sequence effects on preferences for Option I and observed response bias in the third choice task

Sample	Sequences Effects in Option I			Bias in Third Task Responses Compared to Control Case		
	First-order Initial (Observed)[a]	First-order Secondary (Potential)	Second-order (Potential)	Direction (Observed)[a]	One-Tailed Test	Two-Tailed Test
A_1	**+**	+	+	**+**	0.0587	0.1173
A_2	**None**	+	+	**+**	0.0306	0.0612
B_1	**–**	+	None	**None**	0.4274	0.8547
B_2	**+**	**–**[b]	+	**None**	0.1973	0.3947
C_1	**None**	+	+	**+**	0.0587	0.1173
C_2	**+**	None	+	**+**	0.0003	0.0006

Notes:

[a] An observed sequencing effect is classified as being positive (+) or negative (–) if a one-tailed test is significant at the 90% confidence level or greater.

[b] See discussion in footnote 1

FIGURES:

Figure 1: Examples of Price-sequencing in a Choice Experiment

Option II in an Improving Price Sequence:

Choice Task 1	
Option I	**Option II**
Small	Large
€0	**€High**

⇨

Choice Task 2	
Option I	**Option II**
Small	Large
€0	**€Low**

Option II in a Worsening Price Sequence:

Choice Task 1	
Option I	**Option II**
Small	Large
€0	**€Low**

⇨

Choice Task 2	
Option I	**Option II**
Small	Large
€0	**€High**

Figure 2: Examples of Commodity-sequencing in a Choice Experiment

Option I in an Improving Commodity Sequence:

Choice Task 1	
Option I	**Option II**
Small	Large
€0	€High

⇨

Choice Task 2	
Option I	**Option II**
Medium	Large
€0	€High

Option I in a Worsening Commodity Sequence:

Choice Task 1	
Option I	**Option II**
Medium	Large
€0	€High

⇨

Choice Task 2	
Option I	**Option II**
Small	Large
€0	€High

Figure 3: Examples of Mixed Pricing and Commodity-sequencing in a Choice Experiment

Improving Commodity (Option I) and Price (Option II) Sequences:

Choice Task 1	
Option I	**Option II**
Small	Large
€0	**€High**

⇒

Choice Task 2	
Option I	**Option II**
Medium	Large
€0	**€Low**

Worsening Commodity (Option I) and Price (Option II) Sequences:

Choice Task 1	
Option I	**Option II**
Medium	Large
€0	**€Low**

⇒

Choice Task 2	
Option I	**Option II**
Small	Large
€0	**€High**

Figure 4: Health state descriptions used in the study

Severe Ill-Health: 22223

- **I have some problems in walking about**
- **I have some problems with self care**
- **I have some problems in performing my usual activities** (work, study, housework, family or leisure activies)
- **I have moderate pain or discomfort**
- **I am very anxious or depressed**

Mild Ill-Health: 21212

- **I have some problems in walking about**
- **I have no problems with self care**
- **I have some problems in performing my usual activities** (work, study, housework, family or leisure activies)
- **I have no pain or discomfort**
- **I am very anxious or depressed**

Figure 5: Schematics of typical choice task

	Option I	Option II
Duration of Treatment	12 mths	12mths
Symptoms	Severe	None
Duration of Symptoms	2 mths	None
Cost	€0	€120

I	Severe
II	€Med

Figure 6: Experimental design and observed acceptance rates for options in each task

Sample	Choice Task 1			Choice Task 2			Choice Task 3		
A_1	A_{11}			A_{12}			A_{13}		
	I	Severe	25.3	I	Severe	42.2	I	Mild	49.4
	II	€Med	74.7	II	€High	57.8	II	€Med	50.6
A_2	A_{21}			A_{22}			A_{23}		
	I	Severe	32.5	I	Severe	24.1	I	Mild	36.1
	II	€High	67.5	II	€Med	75.9	II	€Low	63.9
B_1	B_{11}			B_{12}			B_{13}		
	I	Mild	37.3	I	Severe	8.4	I	Mild	24.1
	II	€Med	62.7	II	€Med	91.6	II	€Low	75.9
B_2	B_{21}			B_{22}			B_{23}		
	I	Severe	25.9	I	Mild	64.7	I	Severe	38.8
	II	€Med	74.1	II	€Med	35.3	II	€High	61.2
C_1	C_{11}			C_{12}			C_{13}		
	I	Mild	22.9	I	Severe	32.5	I	Mild	49.4
	II	€Low	77.1	II	€High	67.5	II	€Med	50.6
C_2	C_{21}			C_{22}			C_{23}		
	I	Severe	28.9	I	Mild	41.0	I	Severe	59.0
	II	€Med	71.1	II	€Low	59.0	II	€High	41.0

Notes:

Top box of each choice task represents Option I, bottom box Option II.
Acceptance rates are printed in the box to the right of each Option
Severe = 2 months 22223, Mild = 2 months 21212
High = €240, Medium = €120, Low = €60

Figure 7: First- and second-order sequences in a three-task choice experiment