Dundee Discussion Papers in Economics 153

Jones, Martin K.; Seaman, Paul T.

Publication date:
2003

University of Dundee

Dundee Discussion Papers in Economics

Sampling Bias in Economics Experiments: An Analysis of the Effects of Attrition among Subjects who Sign Up for Experiments

Martin K. Jones
and
Paul T. Seaman

Department of
Economic Studies,
University of Dundee,
Dundee.
DD1 4HN

Working Paper
No. 153
October 2003
ISSN:1473-236X

Sampling Bias in Economics Experiments: An analysis of the effects of attrition among subjects who sign up for experiments**

Martin K. Jones * † & Paul T. Seaman *

October 2003

Abstract

Experiments in economics tend to use a restricted subject pool consisting of university students. This obviously raises concerns about the possibility of sampling bias. Although work has been undertaken in particular experiments to compare the results obtained from different subject pools, there has been little systematic investigation of the potential problems caused by the subject recruitment methods used. In this paper we examine one particular aspect of this – the fact that students who sign up for experiments do not always attend. Using two 'classic' experiments we tested for the existence of this form of sampling bias, and our results indicated its presence in one of the two experiments undertaken. This suggests that the issue of sampling bias is an important subject for further research in experimental economics, and an issue to be carefully considered in the design and implementation of experiments.

* Department of Economic Studies, University of Dundee, Dundee, Scotland. DD1 4HN

** The authors would like to acknowledge the financial support of the Economic and Social Research Council (Award R000223606), the help of Qing Lu in recruiting subjects and helping to organize the experiments and the laboratory assistance of both Qing Lu and Lesley Luo.

† Corresponding Author : m.k.jones@dundee.ac.uk

I. Introduction

One of the major criticisms of the use of experiments in economics has been its lack of 'realism' when compared to the world outside the experimental laboratory. This issue is termed 'parallelism'; can it be said that there is any correspondence between how subjects behave in the laboratory and how people behave in the external world? The general response in the literature to this question has been to question whether it is meaningful in its most general sense – one cannot judge 'parallelism' as such, but one can judge specific scenarios to determine whether or not parallelism holds in a particular case. This can be undertaken by isolating particular aspects of the 'real world' and testing these aspects in the laboratory or by using econometric analysis on non- experimental data.

This answer to the question of parallelism has not satisfied everyone on a logical level since such tests may always have a problematic interpretation (see Starmer (1999)). However, this does not mean that methodological tests of this type do not have a major role to play in experimentation as can be seen from previous literature on the subject. Typical examples of such 'methodological testing' include testing the random lottery incentive system (see Cubitt, Starmer & Sugden (1998)), the binary lottery elicitation mechanism (Selten, Sadrieh & Abbink (1995), Loomes (1998)), hypothetical decisions (Beattie & Loomes (1997)) and different stochastic forms for decision functions (Loomes & Sugden (1998)).

One area of methodological testing that is of direct relevance to the issue of parallelism and which has been of interest to many experimentalists has been that of sampling. It is obvious that the sample which is most often used by experimental economists – students selected from the university population – is not representative of the population as a whole. Furthermore, the samples used in experiments are quite often self-selecting. This self-selection takes place on two levels – (1) students volunteer for the experiment and then (2) a subset of these volunteers actually attend the experiment. It is clear, therefore, that the sample used is certainly not a random sample of the university population, never mind the whole population. The 'casual' defence which seems to be most often offered is that while this sampling may not be representative, or even randomly selected, there is nothing about the subjects in the sample that would result in biased results from analysis of their choices. This *a priori* defence is frequently followed by a challenge to critics who might suggest potential biases to test them in the experimental laboratory.

There are many cases where this defence may well not be a sound argument, as well as cases where the subject pool does matter and may cause biased results to emerge from analysis of the choices observed. In many of these cases the potential biases have been tested in the laboratory. One example of this is the ultimatum game where the results typically show consistent evidence of 'fairness' between the two players in the game. Since these players are usually from a similar background, one issue is whether we would see different behaviour if the players were from different backgrounds. Two examples of experiments with different samples include experiments focusing on gender differences (Eckel & Grossman (2001); Solnwick (2001)) and cross-cultural differences (Cameron (1999); Costa-Gomes & Zauner (2001)). In terms of their respective choices these studies frequently found substantial differences between the samples.

Even within individual choice theory, where which types of people are in a sample is not so obviously important, there are many examples of experimenters testing different samples of subjects. Two examples of this are the preference reversal phenomenon – Lichtenstein and Slovic (1973) tested the phenomenon in a Las Vegas Casino – and risk aversion, which Binswanger (1980) tested in an Indian village. In these and other examples it was generally found that there was little difference between behaviour in the laboratory and behaviour in the outside world.

While there has been extensive testing of hypotheses with different samples of subjects, there seems to have been few examples of direct comparisons between different samples. In particular there have been few comparisons that have directly tested the problem of parallelism between university students who volunteer for (and attend) experiments and the rest of the population. It is the aim of this paper to examine part of this problem as a preliminary step to investigating the problem as a whole. The sampling problem this paper is concerned with is that of the attrition of those students who volunteer for experiments. The paper will investigate whether there is any difference between those students who volunteer for the experiment and those who actually attend the experiment. Section II will set out the methods used in this paper and describe the data collected. Section III will present the results of our analysis, focusing on the evidence of sampling bias in those results. Section IV will summarise the results of the paper, and offer suggestions for further research in this area.

II. Methods And Data

Each experimental session consisted of two parts; a questionnaire completed by students as part of the process of volunteering for the experiment (this questionnaire will be described later) and at a later date the actual experiment itself which was only undertaken by those subjects who actually attended the experiment. It was decided that in order to perform tests that would be widely recognised as authoritative then the tasks chosen should be already well-tested in other contexts and produce consistent and robust results. This suggested that the ideal type of experiments were ‘classic’ experiments focusing on individual choice under conditions of risk. The most obvious choices (because of their widespread use and their consistent results) were the tasks used to test for the Common Consequence and Common Ratio effects (Allais (1953)).

The Common Consequence effect was discovered as the result of experiments involving two pairs of choices over lotteries. In the first pair of lotteries the subject is given a choice of two lotteries which are structured in the same way as this example:

Lottery 1: 100% chance of £10
Lottery 2: 5% chance of £20; 90% chance of £10; 5% chance of £0

Then the subject has the choice of a second pair of lotteries as follows:

Lottery 3: 10% chance of £10; 90% chance of £0
Lottery 4: 5% chance of £20; 95% chance of £0.

If the subjects in the experiment follow Savage's (1954) theory of subjective expected utility then they should either choose lottery 1 over lottery 2 and lottery 3 over lottery 4 or choose lottery 2 over lottery 1 and lottery 4 over lottery 3. Any other pattern of choices would violate the independence axiom of expected utility. This is because Lotteries 1 and 2 have a common outcome (90% chance of £10) as do lotteries 3 and 4 (90% chance of £0). If these are ignored (as Savage's independence axiom requires) then the two sets of lotteries are identical.

The Common Ratio effect also involves two pairs of choices over lotteries. In the first pair of lotteries the subject is given a choice of lotteries as follows:

Lottery 1: 80% chance of £10; 20% chance of £0
Lottery 2: 40% chance of £20; 60% chance of £0

Then the subject has the choice of lotteries as follows:

Lottery 3: 10% chance of £10; 90% chance of £0
Lottery 4: 5% chance of £20; 95% chance of £0

In this case subjects should again either choose lottery 1 over lottery 2 and lottery 3 over lottery 4 or choose lottery 2 over lottery 1 and lottery 4 over lottery 3. Again, any other pattern of preferences violates the independence axiom of expected utility. In this case this is because the independence axiom suggests that if the chances of winning positive amounts of money are multiplied by a constant then the same preference ordering should be preserved.

In the experimental sessions described in this paper the pairs of lotteries outlined above were those used in our experiments. In all six choices were given between pairs of lotteries within which the four lotteries outlined above were incorporated. Each choice between a pair of lotteries constituted a question in the experiment These six questions were given in random order for each person. It was decided to give the lotteries a visual representation in a form that has frequently been used before (Appendix 4 presents the graphical representation for one of the six questions we used).

During each experimental session the subjects who attended were given seats in the experimental laboratory. These subjects were unable to see each others' answers to the questions asked in the session and were told not to communicate in any way. The experimenter then went through a series of instructions explaining first of all the lottery diagrams and how they would be played out and then how to make a choice between two lotteries. The subjects then went through a practice question to see if they understood the experiment – this practice question simply tested for dominance and all but seven of the 179 subjects answered this question correctly. At several points during the instructions the subjects were given the opportunity to ask questions if they did not understand something.

Once the experimenter was sure that everyone understood the experiment the subjects were then told to answer all six questions in their own time. When everyone had finished the experiment the experimenter then went round each subject in turn, rolling a six-sided die to select which one of the six questions would be played out for real. The lottery chosen in this question was played out using a 20-sided die and the subject received the amount of money which they won in that lottery[1].

The nature of the experiment we wanted to perform required information not only on those students who signed up for an experiment and attended the experiment to which they were allocated, but also those students who signed up for an experiment, were allocated an experiment to attend, but failed to attend. Thus, we needed as much information as possible from all students who signed up for an experiment at the time they signed up. It was decided that the most suitable method for achieving this would be a web form using (where possible) drop-down menus, since it offered the following advantages over the alternative methods:-

- it would be less costly than providing paper forms to all students who *potentially* might wish to sign up for an experiment

- it would result in fewer Data Protection Act problems than asking students to complete their information on a poster on a noticeboard

- it would result in fewer students providing incomplete / inaccurate information than asking them to sign up for an experiment via a normal (unstructured) email

- It was by far the easiest method for students to fill in the questionnaire and required the minimum of effort

The main method of recruiting students was by email (with a link in the email to the URL of the web form). Appendix 5 depicts the content and structure of the web form used.

[1] Students were paid £2 simply for attending the experiment, and the average payout to students from the experiment itself was around £8, making a total average payout of around £10.

The design / content of the web form presented us with several dilemmas, the most important of which was how much information to ask of students when signing up – too little information would restrict our empirical analysis, too much information would discourage students from signing up. Furthermore, once we had decided broadly how much information we should gather, there was little by way of formal economic theory that we could use to decide what specific questions to ask students when they signed up. Ultimately our decision regarding what information to ask of students was decided on the basis of what seemed 'sensible' with reference to 'common sense'. After careful consideration we decided to obtain the following information (in addition to their name and email address):-

- personal information – gender, age (in bands – 18-24, 25-44, 45 and over), marital status, whether English is their first language

- university information – faculty, department(s)[2], year of study, whether lives in university halls of residence (and if so, which one)

- timetable preference – their first, second and third preference amongst a choice of up to eight experiment times[3]

This information, explicitly obtained from students, was augmented with the date they signed up for an experiment, and the local weather conditions on the day of the experiment to which they were allocated (information on rainfall, hours of sunshine and whether snow fell or lay on the day in question was obtained for a local weather centre from the UK Meteorological Office). Combining information together meant that other variables could be generated (e.g. knowing when the student signed up and when their allocated experiment was meant that we could calculate how long they had to wait before attending their allocated experiment).

[2] Because of the large number of departments within the University of Dundee the department(s) information was the only item out of these three groups of information that was not obtained from a list on the web form, but rather typed in by students. The vagaries of the responses from students (including numerous cases of no response or ambiguous response) meant that this information had to be discarded.

[3] Every student was allocated an experiment time that matched either their first, second or third preference. There was 54% attendance for those allocated an experiment time matching their first preference, compared to 46% for those with an experiment time matching their second or third preference. We deliberately ensured that through the majority of students were allocated to their first preference time, a significant minority were allocated to their second or third preference.

The inclusion of ‘personal information’ in the available data reflects accepted practice in other areas of economics (e.g. labour economics) but is also relevant given ‘perceived wisdom’ – e.g. females often have a better attendance and performance record in class, older students often have difficulty adjusting to academic study, married students (particularly those with children) are more likely to have to respond to ‘family events’ and therefore miss their allocated experiment, and students with a less clear grasp of English may have difficulty in truly understanding the instructions given to them.

The inclusion of ‘university information’ in the available data reflects the fact that the nature of a student’s studies may affect whether they attend or how they perform – e.g. the lotteries in the experiment involve numerical skills and students from some subject areas (the humanities, art, design, etc.) may feel less comfortable with such calculations, students in later years of study may have a better skill base or the confidence to try something out of the ordinary (like experiments), while living in particular forms of university halls of residence may affect signing up and attendance (students may be more likely to attend if their flatmates are due to attend at the same time) or even performance (students may learn from the experience of their friends who have already taken the experiment).

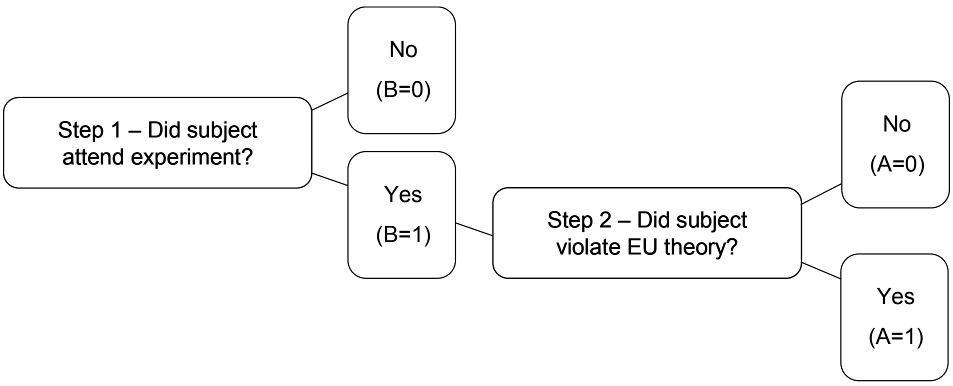
The inclusion of ‘timetable preference’ in the available data permits us to examine whether the ‘ease’ with which students can attend is a factor in attendance and performance – it may be the case that students are unwilling to ‘go out of their way’ to attend their allocated experiment, and those that do are perhaps more determined to reap the potential benefits from the experiment and therefore perform better.

In analysing the results of this experiment we work in a stochastic choice framework which is applied to each subject for their decisions to attend and, if they did attend, whether or not they violated the independence axiom in the experiment. Stochastic variation in choice within the experimental session comes from errors or imprecisions in the individual's preferences and beliefs (Loomes & Sugden (1995)). Stochastic variation in the decision to attend the experiment derives from errors by the subjects (e.g. forgetting the time of the experiment / timetable clashes / not knowing where the experiment was going to be held) or from unforeseen circumstances (e.g. illness / personal problems). It is assumed that these are essentially random circumstances and do not apply to any particular groups of people[4]. By modelling choice as stochastic we allow for the possibility of subjects' behaviour being partly random. Our econometric model tries to establish non-random elements in that behaviour.

Each effect (Common Consequence and Common Ratio) is the result of choices in two lotteries. It is assumed that these two choices can be represented by one random variable *A* which is equal to 0 if the subject followed expected utility theory and 1 otherwise. We assume that *A* has a standard normal distribution. While the assumption of normality could be seen as arbitrary (although one could appeal to the Central Limit Theorem to support it), similar assumptions have been made elsewhere in the literature (Hey & Orme (1994), Loomes, Moffat & Sugden (2002)) and no convincing alternative has been put forward. We also assume that whether a person attends the experiment is a random variable *B* which is equal to 1 if the subject turned up to their experiment and 0 if they did not, and it is likewise assumed to have a standard normal distribution. When estimating variables *A* and *B* jointly we assume they vary with a bivariate normal distribution.

The diagram below illustrates the two step model which will be estimated in this paper[5].

Figure 1



These two steps could be estimated separately, and standard probit estimation would seem a perfectly sensible way in which to proceed in that case. However, one of the issues we

[4] It could be argued that some of these could be non-random e.g. a virus could spread rapidly round a hall of residence but not affect another hall. However no such event took place during the time of the experiment.
[5] There are actually two models, one for testing the common consequence effect and the other for testing the common ratio effect, but apart from the specifics of the effect being tested the two models are identical in structure.

seek to examine in this paper is whether or not it is appropriate to estimate the two steps of our model separately; given that we observe student choices only for those students who actually attend their allotted experiment session we have censored data, and hence the model we choose to estimate is a bivariate probit model with sample selection[6]. This model permits us to take account of the possibility that the student's decision to attend their allocated experiment is correlated with the decisions they take within the experiment itself if they do attend.

III. Results

The experiments were carried out during the first four months of 2003 on students at the University of Dundee (these students were almost exclusively full-time undergraduates). Excluding the small number of students failing to submit all the required information on the web form (a requirement to be included in our sample) left us with a working sample of 340 students who signed up for an experiment. Of these 340 students, 177 of them (i.e. 52%) attended the experiment to which they were allocated.

Appendix 1 presents descriptive statistics for our three distinct groups of students – (a) all students in our sample, (b) those students who attended their experiment, and (c) those students who did not attend their experiment. These data reveal sizeable differences in the characteristics of the different groups; for example, whereas Third Year students constitute 33.3% of attending students, they constitute only 19% of non-attending students, and while students living in Belmont or Springfield Halls constitute 15.3% of attending students, they constitute only 2.5% of non-attending students. Therefore, it is evident that the sub-sample of students who attended their experiment is not representative of the sample of students who signed up for the experiment via the web form, and as a result sample selection bias is a possibility.

As previously indicated, the choices students made would indicate whether or not they diverged from Expected Utility through the Common Consequence and Common Ratio effects. Of the 177 students who attended their allocated experiment there were 48 students (27%) whose decisions were consistent with the Common Consequence effect and 88

[6] For a full description of this model see Greene (2003).

students (50%) whose decisions were consistent with the Common Ratio effect. Table 1 below summarises the outcomes of these experiments.

Table 1 : Student performance in the experiments			
	Common Ratio Effect Absent	Common Ratio Effect Present	Total
Common Consequence Effect Absent	66	23	89
Common Consequence Effect Present	63	25	88
Total	129	48	177

One aspect of interest in Table 1 is the fact that the Common Consequence and Common Ratio effects seem to be completely independent of each other. While this is unusual, it can be explained in terms of the common non-expected utility theories of choice. If we assume that subjects have preferences which 'fan out' in the sense of Machina (1982) then, since the lotteries for testing the Common Ratio and Common Consequence effects have different slopes in the Marschak-Machina triangle, one would expect subjects with different 'fannings out' of indifference lines to give different answers to the questions. Nevertheless, it is of interest that the two effects should be independent in this manner.

Appendix 2 presents the bivariate probit results for both the Common Consequence effect and the Common Ratio effect. As previously indicated, the specifications for the two models tested are identical apart from the second step dependent variable. Furthermore, for both models identification is achieved through non-linearity in the bivariate probit model and by the inclusion in the first step of three variables likely to affect attendance but not performance – weather conditions (proxied by that day's rainfall), whether or not the student was allocated to an experiment matching their first time preference, and the number of days between them filling out the web form and the day of their allocated experiment.

The first point to be taken from Appendix 2 is that for the Common Ratio effect the χ^2 statistic of 4.11 ensures that we can reject at the five percent level the null hypothesis that the two steps can be estimated independently; however, for the Common Consequence test the χ^2 statistic of 1.51 is considerably short of the level required to make a similar conclusion. The conclusion to be drawn from this is that the results obtained from experiments *may* be subject to sample selection bias; however, it is not immediately obvious *a priori* which specific experiments will generate results suffering from sample selection bias.

We now turn to the three variables included in the first step only. For the variable indicating that the student was allocated to an experiment time matching their first time preference we find a positive and significant effect in the Common Consequence test only. One reason why we did not obtain a stronger / more consistent result could be due to the fact that all students were asked to state their three most preferred time slots out of either seven or eight time slots that differed in terms of either time or date, and all students were allocated one of the three slots they chose. Had we allocated some students to one of the four or five time slots they did not list as one of their three preferred time slots then the degree of difficulty in attending might well have been heightened for those students and therefore more of an effect found. However, the results as found do indicate that the process by which students are allocated to specific experiments may affect attendance at the experiment.

The weather conditions on the day of the experiment were accounted for by the inclusion of the inches of rain that fell on the day in question (an alternative specification, incorporating hours of sunshine rather than inches of rain, was estimated but added nothing to the results, and hence those results are not presented in this paper). Somewhat bizarrely higher levels of rainfall improved attendance in the Common Ratio test (though only at the ten percent level) though it had no effect whatsoever in the Common Consequence test. However, we can still argue that weather conditions did little to *significantly* affect the attendance patterns that we observed.

The variable capturing the number of days between when a student submitted their information via the web form, and when the experiment they were allocated to was held, was not significant. However, the fact that the 'advertising' of the experiments was targeted at different groups of students at different times during the semester, and students were always allocated to an experiment within the subsequent couple of weeks or so, ensured that the gap was never long for any student, and so it is not perhaps surprising that no significant effect was found. The effect on attendance of student memory (and organisation / time-management skills) could be revealed by (a) increasing the delay for some students and (b) sending out reminder emails to some students but not to other students.

There is only very weak support for our prior belief that females may be more likely to attend their allocated experiment, but we do find that females are significantly less likely to fail the Common Ratio effect (for the Common Consequence effect the female effect is just short of significance at the ten percent level). Speaking English as one's native language had no discernable effect, which could be because of (a) the experiments being relatively simple by the standards of the experimental economics field, (b) the instructions being set out in a particularly clear, logical manner, or (c) the English language induction programme put on by the University of Dundee for new students for whom English is not their native language[7]. Neither did we find any effect for the student being a mature student (i.e. 25 years or older).

Students from the Faculty of Law and Accountancy were more likely to attend their allocated experiments (compared to the default faculty category of Arts and Social Sciences), though once again there is no effect on performance. On the other hand students from the Faculty of Duncan of Jordonstone (primarily an arts and design college, but incorporating architecture) were, at the ten percent level, more likely to fail the Common Ratio test; many of these students will have relatively few opportunities to develop their numerical skills. In terms of the year that a student was from, we find that relative to the Year 1 default the most reliable attendees were Year 3 students, and these students also performed better in the experiments (though for the Common Ratio effect test the coefficient is significant at only the ten percent level); one explanation for this result is that at least some of these students may have had lower academic burdens than other students – the ‘standard’ degree length at Scottish universities is 4 years, and at the time there were still a number of degrees at the University of Dundee with undivided honours exam diets – all third and fourth year exams would be taken at the end of the fourth year.

The University of Dundee has ten main halls of residence for student accommodation, and we found that three yielded significant results – students living in either Belmont or Springfield Halls (large complexes close to the location of the experiments that offer cheaper accommodation for typically new, first year students) were more likely to attend and less likely to fail the Common Ratio effect, while students living in West Park Halls (a smaller complex further from the location of the experiments that offers more expensive accommodation that is generally reserved for older and foreign students) were less likely to

[7] It was also the case that by the time students attended their allocated experiment they would have been living in Dundee and attending English language courses for at least four months and so would have improved their level of English language comprehension considerably.

attend. Thus, attendance and performance were superior for those students for whom it was (a) easier to attend, (b) more likely they would interact with other students who had signed up for and / or already attended their experiment, and (c) likely to result in a monetary return higher than they could obtain from other income generating activities (assuming that first year students are on the lowest rung of the undergraduate job ladder).

One final variable that we included in our specification concerns the failure of students to submit time preference information as per the clear instructions included in the web form. Students were to provide a first preference, a second preference and a third preference, and these three preferences were not to relate to the same time slot (for each of these three preferences, asked one after the other, the choice was made by selecting from a 'drop down list' of the same set of eight time slots – two slots on each of four days (Monday, Tuesday, Wednesday and Friday)). The specification includes a (0,1) dummy variable for whether the student made a time preference data error – either failing to give three valid time preferences or using the same time slot for more than one time preference. We interpret such 'failings' as indicative of a failure to comprehend clear instructions or a lack of attention to detail, either of which may result in being less likely to attend the experiment they were allocated to and / or (should they attend) more likely to fail one or both of the two effects being tested. The results strongly support the *a priori* beliefs (though for the Common Consequence test the result is significant at only the 10% level).

We generally followed the rule of eliminating from our specifications those variables which were consistently found to be statistically-insignificant (bearing in mind that there is no formal economic theory determining what *should* be included as regressors). In particular, we found no evidence of the following variables having any significant effect – the marital status of the student, the day of the week or the time of the day of the experiment they were allocated to, or a variety of interaction terms such as those between gender, age and marital status. In none of these various specifications were these variables even close to being statistically significant, and therefore were excluded from the results presented in this paper.

In the early part of this section we noted that for the Common Ratio effect only a likelihood ratio test indicated that we could reject at the five percent level the null hypothesis that the two steps can be estimated independently. In other words, estimating the second stage without the first (selection / attendance) stage would dramatically change the results obtained.

The extent of this change can be gauged from Appendix 3, which presents for the second stage only results obtained (a) with and (b) without joint estimation with the first (selection / attendance) stage (the second set of results are those obtained from a simple probit estimation). For the Common Consequence effect we find that the constant and one variable were (in the absence of joint estimation) insignificant, but with joint estimation are now significant at the 5% level. For the Common Ratio effect we find that the effect of joint estimation is that the constant ceases to be significant at the 5% level and two variables become significant at the 5% level. As one would expect, changes in the significance of variables are associated with changes in the coefficients of variables (with some coefficients doubling or more, or halving, as a result of joint estimation). Although there are changes in the results of both models, the greatest changes are seen in the Common Ratio effect results; this is as we would expect given the likelihood ratio tests discussed earlier, where joint estimation was shown to be appropriate for the Common Ratio effect.

To conclude this section we can state that our results strongly support the argument that the correlation between the student's decision to attend an experiment and their performance in the experiment should they attend is such that joint estimation, such as that provided by the bivariate probit model, is appropriate. Furthermore, the results obtained from such an estimation procedure on the data we collected produced results that in many cases support our *a priori* beliefs – in particular, 'personal', 'university' and 'timetable preference' information constitute useful predictors of individual student behaviour in terms of both stages in our two-stage model.

IV. Conclusions

The aim of this paper was to discover whether the fact that subjects who volunteered for an experiment didn't always attend caused a bias in the results of the experiment. We found evidence of attrition-based sampling bias in our results for the Common Ratio effect but not in our results for the Common Consequence effect. This implies that there is a qualitative difference between the two experiments, in particular that the Common Ratio effect is much more susceptible to vary in magnitude as a result of different types of people in the sample.

This is illustrated by the independent variables used in the regression. In general, the variables which are significant for those subjects who violated Expected Utility theory in the Common Ratio test are those which suggest either a weakness in mathematical ability or a lack of attention to detail. Meanwhile, those who tended not to violate Expected Utility theory often had an incentive or the time to pay attention to the task. This is not so much the case for the test for the Common Consequence effect where only one variable is significant (for third year students). This suggests that the Common Consequence effect is probably not particularly susceptible to which types of people are in the sample. By contrast, at least part of the cause of the Common Ratio effect may be traced to different types of people in the sample[8].

The obvious question in this situation is: why should this be the case? Our argument is based around the idea that the Common Ratio test is more 'intuitive' than the Common Consequence test. In the Common Ratio test one pair of lotteries is simply a scaled down version of another pair, so this can be spotted comparatively easily by someone paying attention. This 'proportionality' heuristic, we claim, is comparatively easy to apply and so it is relatively easy to be consistent in this case. The results of the analysis suggest that certain types of people pay attention more than others and this enables them to conform more closely to what is predicted by expected utility theory. By contrast, the Common Consequence test requires the subject to notice that the two pairs of lotteries are identical apart from their two common consequences. This may not be spotted (or even seen as relevant if they are spotted). This means that it would be more likely that the Common Ratio test would be more influenced by what sort of people turn up to the experiment or not[9].

On a more general level it seems that there is a problem of attrition in experiments as a result of subjects not attending sessions for which they have signed up. While this does not apply to all experiments, it does suggest that there is a certain class of experiments which may be affected by people who do or do not attend their experiment. Furthermore, the link seems to be, in general, a negative one in that subjects from certain groups who are less likely to turn up are more likely to violate Expected Utility and vice versa. This suggests that experiments in these cases may *underestimate* the levels of violations of Expected Utility

[8] In this context it is interesting that the two effects seem to be independent of each other.

[9] We do not want to claim here that either the Common Ratio or Common Consequence effects are purely perceptual biases (which in turn assumes that there is a 'correct' way of deconstructing the lotteries). We are simply claiming that the proportionality aspect of the Common Ratio effect does appeal on grounds of generality and simplicity as a heuristic more than common consequences do in the Common Consequence effect.

involved over the whole population of those who actually sign up for an experiment.

This in itself is ‘good news’ for those experimentalists who would like to see the violations of Expected Utility made more robust. However, it does mean that experimentalists can no longer ignore the environment in which they work. Sampling issues are as important in experiments as in any other type of empirical research and they need to be investigated thoroughly to ensure that experimental economics results derived from laboratory experiments are as robust as those coming from other types of data and from other fields of economics.

Bibliography

Allais M. (1953) "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de L'Ecole Americaine" *Econometrica* vol. 21 p. 503 - 546

Beattie J. and Loomes G. (1997) "The impact of incentives upon risky choice experiments" *Journal of Risk and Uncertainty* vol. 14 p. 149- 162

Binswanger H.P. (1980) "Attitudes towards Risk: Experimental Measurement in Rural India" *American Journal of Agricultural Economics* vol. 62 p. 395- 407

Cameron L.A. (1999) "Raising the Stakes in the Ultimatum Game: Experimental evidence from Indonesia" *Economic Inquiry* vol. 37 p. 47-59

Costa- Gomes M. and Zauner K.E. (2001) "Ultimatum bargaining Behavior in Israel, Japan, Slovenia and the United States: A Social Utility Analysis" *Games and Economic Behavior* vol. 34 p. 238-269

Cubitt R., Starmer C. and Sugden R. (1998) "On the validity of the Random Lottery Incentive Mechanism" *Journal of Experimental Economics* vol. 1 p. 115-130

Eckel C.C. and Grossman P.J. (2001) "Chivalry and Solidarity in Ultimatum Games" *Economic Inquiry* vol. 39 p. 171- 188

Greene W.H. (2003), *Econometric Analysis* (fifth edition), Prentice Hall

Hey J.D. and Orme C. (1994) "Investigating Generalizations of Expected Utility Theory Using Experimental Data" *Econometrica* vol. 62 p. 1291- 1326

Lichtenstein S. and Slovic P. (1973) "Response - induced reversals of preference in gambling: an extended replication in Las Vegas" *Journal of Experimental Psychology* vol. 101 p. 16 -20

Loomes G. (1998) "Probabilities vs Money: A test of some of the fundamental assumptions about rational decision making" *The Economic Journal* vol. 108 p. 477-89

Loomes G. and Sugden R. (1995) "Incorporating a Stochastic Element into Decision Theories" *European Economic Review* vol. 39 p. 641-648

Loomes G. and Sugden R. (1998) "Testing different stochastic specifications of risky choice" *Economica* vol. 65 p. 581- 598

Loomes G., Moffatt P. and Sugden R. (2002) "A Microeconometric Test of alternative Stochastic Theories of Choice" *Journal of Risk and Uncertainty* vol. 24 p. 103-130

Machina M. (1982) "'Expected Utility' theory without the independence axiom" *Econometrica* vol. 50 p. 277-323

Selten R., Sadrieh A. and Abbink K. (1990) "Money does not induce risk neutral behavior, but binary lotteries do even worse" *Theory and Decision* vol. 46 p. 211-249

Solnwick S.J. (2001) "Gender Differences in the Ultimatum Game" *Economic Inquiry* vol. 39 p. 159 - 200

Starmer C. (1999) "Experiments in Economics….. (should we trust the dismal scientists in white coats?)" *Journal of Economic Methodology* vol. 6 p. 1-30

Appendix 1 : Descriptive Statistics

Variable	All subjects	Attending subjects	Non-attending subjects
Student is female	0.497	0.514	0.479
English is student's native language	0.841	0.797	0.890
Student is a mature student	0.091	0.102	0.080
Student is in Faculty of Law & Accountancy	0.056	0.096	0.012
Student is in Faculty of Duncan of J'stone	0.112	0.085	0.121
Student is in Faculty of Engineering	0.176	0.186	0.166
Student is in Faculty of Life Sciences	0.256	0.266	0.245
Student was in Year 2	0.191	0.158	0.227
Student was in Year 3	0.265	0.333	0.190
Student was in Year 4, 5 or postgraduate	0.200	0.192	0.209
Student made time preference data error	0.156	0.119	0.196
Student lived in Belmont / Springfield Halls	0.091	0.153	0.025
Student lived in West Park Halls	0.068	0.028	0.110
Number of observations	340	177	163

Appendix 2 : Empirical Results From Joint Estimation

Variable	Common Consequence Effect		Common Ratio Effect	
	Coeff	t-stat	Coeff	t-stat
Attended experiment (B) = 1				
Student is female	0.222	1.46	0.242	1.61
English is student's native language	0.011	0.05	-0.315	-1.41
Student is a mature student	0.126	0.46	0.222	0.83
Student is in Faculty of Law & Accountancy	1.214	2.74	1.033	2.42
Student is in Faculty of Duncan of J'stone	-0.143	-0.54	-0.246	-0.90
Student is in Faculty of Engineering	0.083	0.39	0.046	0.21
Student is in Faculty of Life Sciences	0.117	0.61	0.178	0.92
Student was in Year 2	-0.055	-0.27	-0.029	-0.14
Student was in Year 3	0.530	2.65	0.531	2.69
Student was in Year 4, 5 or postgraduate	0.079	0.38	0.044	0.21
Student made time preference data error	-0.426	-2.08	-0.509	-2.38
Student lived in Belmont / Springfield Halls	1.127	3.44	1.118	3.47
Student lived in West Park Halls	-0.636	-1.98	-0.586	-1.82
Rainfall on day of experiment	0.0002	0.01	0.030	1.82
Student's first time preference achieved	0.342	1.99	-0.035	-0.24
Gap between first response and experiment	0.006	0.26	-0.004	-0.13
Constant	-0.615	-1.64	-0.014	-0.04
Violated Expected Utility (A) = 1				
Student is female	-0.252	-1.54	-0.443	-2.79
English is student's native language	0.064	0.29	0.158	0.71
Student is a mature student	-0.157	-0.58	0.021	0.08
Student is in Faculty of Law & Accountancy	-0.427	-1.23	-0.190	-0.55
Student is in Faculty of Duncan of J'stone	0.209	0.75	0.483	1.78
Student is in Faculty of Engineering	-0.050	-0.22	0.028	0.13
Student is in Faculty of Life Sciences	-0.095	-0.48	0.078	0.40
Student was in Year 2	0.262	1.10	-0.030	-0.13
Student was in Year 3	-0.552	-2.69	-0.356	-1.79
Student was in Year 4, 5 or postgraduate	-0.215	-0.97	-0.098	-0.45
Student made time preference data error	0.381	1.63	0.515	2.17
Student lived in Belmont / Springfield Halls	-0.279	-1.08	-0.944	-3.48
Student lived in West Park Halls	0.482	1.18	0.516	1.28
Constant	0.877	2.78	0.456	1.52
rho	coeff=-0.998, std err=0.135		coeff=-1, std err=0	
Log likelihood	-325.5485		-300.8641	
Number of observations	340		340	

A χ^2 statistic of 1.51 for the Common Consequence test means that we cannot reject at the five percent level the null hypothesis that the two steps can be estimated independently.

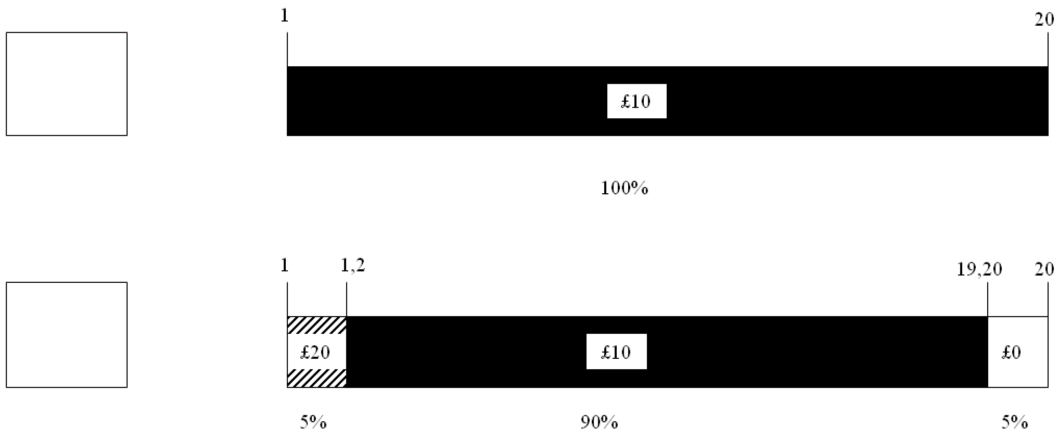
A χ^2 statistic of 4.11 for the Common Ratio test means that we can reject at the five percent level the null hypothesis that the two steps can be estimated independently.

Appendix 3 : Comparison Of Empirical Results From Joint And Independent Estimation

Variable	With selection		Without selection	
	Coeff	t-stat	Coeff	t-stat
Common Consequence Effect				
Student is female	-0.252	-1.54	-0.168	-0.81
English is student's native language	0.064	0.29	0.001	0.01
Student is a mature student	-0.157	-0.58	-0.061	-0.18
Student is in Faculty of Law & Accountancy	-0.427	-1.23	0.029	0.08
Student is in Faculty of Duncan of J'stone	0.209	0.75	0.094	0.25
Student is in Faculty of Engineering	-0.050	-0.22	-0.017	-0.06
Student is in Faculty of Life Sciences	-0.095	-0.48	-0.012	-0.05
Student was in Year 2	0.262	1.10	0.294	0.96
Student was in Year 3	-0.552	-2.69	-0.369	-1.44
Student was in Year 4, 5 or postgraduate	-0.215	-0.97	-0.211	-0.74
Student made time preference data error	0.381	1.63	0.159	0.51
Student lived in Belmont / Springfield Halls	-0.279	-1.08	0.181	0.64
Student lived in West Park Halls	0.482	1.18	0.154	0.25
Constant	0.877	2.78	0.145	0.36
Common Ratio Effect				
Student is female	-0.443	-2.79	-0.511	-2.23
English is student's native language	0.158	0.71	0.299	0.98
Student is a mature student	0.021	0.08	0.212	0.59
Student is in Faculty of Law & Accountancy	-0.190	-0.55	0.748	1.79
Student is in Faculty of Duncan of J'stone	0.483	1.78	0.704	1.74
Student is in Faculty of Engineering	0.028	0.13	0.238	0.76
Student is in Faculty of Life Sciences	0.078	0.40	0.492	1.76
Student was in Year 2	-0.030	-0.13	-0.200	-0.59
Student was in Year 3	-0.356	-1.79	-0.118	0.43
Student was in Year 4, 5 or postgraduate	-0.098	-0.45	-0.089	-0.29
Student made time preference data error	0.515	2.17	0.475	1.42
Student lived in Belmont / Springfield Halls	-0.944	-3.48	-0.443	-1.31
Student lived in West Park Halls	0.516	1.28	0.265	0.44
Constant	0.456	1.52	0.955	-2.19
Number of observations	340		177	

Appendix 4 : Example Of A Question In The Experiment

Below are two lotteries. Please put a cross against the lottery which you would prefer to play out. If this question is picked at the end of the experiment then the lottery which you have chosen will be played out for real.



Appendix 5 : Survey Form For Attrition Experiment

The Experimental Economics Group @ The University of Dundee

PLEASE complete ALL of the questions before submitting your response.

Some questions about you

What is your full name (forename & surname)?	<Text entry box>		
What gender are you?	○ Female	○ Male	
What age are you?	○ 18-24	○ 25-44	○ 45 and over
What is your marital situation?	○ Single	○ Married or living together	
Is English your native language?	○ Yes	○ No	
Do you live in a university residence?	<Drop down list>		

Some questions about your studies

Which Faculty are you in?	<Drop down list>
Which is (are) your departments(s)? [List all]	<Text entry box>
What is your year of study?	<Drop down list>

What would be your preferred times?

My first preference would be	<Drop down list>
My second preference would be	<Drop down list>
My third preference would be	<Drop down list>

And finally your email address

Enter your email address (accurately, please)	<Text entry box>

Once you have completed ***ALL*** of the questions above, please click on the **SUBMIT** button below to sign up for the experiment. You will be notified shortly by email of the time and location to attend. Please attend ***AT THAT TIME ONLY***.

Submit

The Experimental Economics Group collects this information for academic research purposes only. The data will not be used for commercial purposes. Furthermore, we guarantee that data will only be provided to others for the purpose of academic study, and any data that is provided will be given in an anonymous form.