Seminar 2 - Methodology & Design

This is a series of short-answer questions designed to check your understanding of experimental economics in practice.

1. **You think X is the reason behind a particular form of behaviour and are considering how to isolate that effect in a laboratory setting. Outline how you could do so using a design by subtraction.**

What is design by subtraction?

If you think X is the reason for behaviour, you would first design an experiment in which X is impossible, but all other explanations are possible. This would then be compared to an experiment where X is possible. This can isolate the effect you want. An important point here is that X should be only thing which is restricted by the difference between the designs. Otherwise, there may be other factors at play. Also, X should not be correlated with any characteristics which may determine behaviour, otherwise it may not be clear what is the actual mechanism (e.g. gender effects on behaviour may be biological, cultural or may indicate wage differences).

Is there a big difference between this and design by manipulation?

In my opinion, there is not really a big difference: the general difference with design by manipulation is the direction with which we are manipulation the trait (positive vs negative); actually the same thing. (It depends on what you consider the baseline)

Example(s) of design by subtraction

In economics an example might be making all incentives the same to see if incentives really change behaviour in different settings.

Another example may be about reputation building influencing say behaviour in a "repeated" one-shot trust game with stranger matching (with full information). One can then completely remove reputation building by making things anonymous and not providing feedback (removing information feedback). Comparisons can then be made to a scenario with full information disclosure on others' interactions.

OR Suppose the hypothesis is that guilt aversion raises reciprocity, with beliefs about others' expectations as the mechanism, one can remove this element by creating a game where expectations are a fixed factor; perhaps providing the objective distribution of others' expectations.

For example, this is the basis of the most famous of natural experiments - twin studies - in which genetics as a rationale is removed by using genetically identical individuals. The good thing is that identical twins occur at random, so we do not lose any generality by only focusing on them (they are somewhat representative of the population). Actually, this example here is usually to rule out the effects of genetics, but not to find the effects of genetics: for example effects of education on income within identical twin have to be non-genetic effects. This would be related to arguments about nature vs nurture in say IQ.

Other uses?

It can also useful in checking the mechanisms by which an effect occurs. Design by subtraction can be used to rule out behaviour which we do not want to influence our results: it is like controlling for an omitted variable.

Relation to Block designs for fixed personal characteristics X

If you feel that fixed factors X like gender may influence treatment effects, then it may be useful to randomise the treatments within these blocks. This ensures that our treatments will be randomised evenly within gender blocks which may help improve statistical power. This also means that for overall treatment effects, we can be even more sure that it is not due to some imbalance between the control and treatment which is influencing results.

2. **You are considering using to use a within-subjects design or between-subjects design. Briefly outline the advantages and disadvantages of each.**

What is a within and between treatment?

Each subject plays all the treatments in a within-treatment design, while each treatment is assigned to random partitions of the sample in a between-treatment design.

Advantage of within-treatment relative to between-treatment

In a within-subject design, each subject is his/her own control. (Could think of it as being like a twin design treatment) This is great as we need not worry about having different characteristics of participants in each treatment (so it is often easier to get significance). This is because your sample is balanced perfectly on all dimensions and it effectively doubles your sample size.

In a between-subject design, we are relying on the law of large numbers and randomisation in order for effects of the traits of the subjects to be "cancelled out". Thus, we need a larger sample size.

Disadvantage of within-treatment relative to between-treatment

However, on the flip side there is the disadvantage of order effects or even fatigue. Order effects are when one treatment affects results in the subsequent treatment: this may be due to learning (real effort games?), priming of certain behaviour (cooperative vs competitive?), relationship building (?), past experience with others.

Controlling for order effects using experimental methods

For example, past experience may be controlled by only providing information on what happened at the end of the experiment in conjunction with using the strategy method[1]. Relationship building can be controlled by anonymity. Rest periods in between treatments to

[1] In the strategy method, the participant (second mover) is asked to make a decision for every possible action/state of a first mover. This allows the second mover's decision to be applied even without discovering the first mover did.

control for fatigue? However, difficult to control directly for "anchoring or priming effects" of a treatment. These may be controlled statistically ex-post following a randomisation procedure.

Controlling for order effects using statistical methods

After randomising the order of between individual treatments, one can include dummies for the order to capture these effects. Some other active controls may be like including period fixed effects.

Choosing between the two?

If order effects are expected to be big, or difficult to control for via the design of the experiment (for example if we are looking at behaviour over time in multiple treatments: for example: public goods game with punishment versus none), a between subject treatment might be more convincing but may require a larger sample size.

3. **You are considering running a multiple round experiment using the same subjects to play the same game multiple times, to gather as much data as possible. What might cause you to reconsider?**

This is an issue which is related to repeated games.

Matching participants when there are strategic interactions

Note that if one is playing a game with the same partner (say prisoners dilemma) then the setting is totally different than a one-shot game as proven in game theory. By having a stranger matching, one can control for this. But we will then face similar issues as when doing a between subject design. Differing past experiences from facing different individuals when interacting can introduce noise.

Learning

The primary problem is the potentially unanticipated effect of learning. In some cases, this is fine (you might actually be trying to test for the importance of learning or it takes some time for participants to reach equilibrium) but otherwise it can confound your results. The lecture notes talks about how you can control for this by training and testing participants before hand.

One example is if we are using real-effort experiments (e.g doing sums). This would then lead to some noise in looking at the effects of say some different payment scheme especially if there is differential learning based on the incentive scheme. Comparing a treatment with incentives versus one with none would include the effect of greater learning in one treatment as well as the pure incentive effect.

Another issue might be fatigue. After making many rounds of decisions, participants might stop thinking: same issues for surveys. This could introduce more noise into the experiment as subjects start choosing randomly?

Incentivising Participants

You would also need to think about how to incentivize subjects across multiple rounds. For example, one may encounter income effects from earning in earlier rounds. This is related to the issue of pay once or many times in the lecture notes. The general norm here is to randomly select rounds for payment: participants are told that this will occur and are only informed of which rounds are selected at the end of the experiment. Some behavioural biases which may occur: risk aversion changes behaviour?

4. **When running a laboratory experiment that involves strategic interaction how does a "strangers" design differ from a "perfect strangers" design?**

Note that this is in the context of a multiple rounds game and usually implies anonymity.

In a "strangers" design subjects are randomly assigned to play against the other subjects in the session. Under "perfect strangers" this randomization is modified to ensure that subjects never play the same person twice.

Usually perfect strangers is ideal, but is difficult when you are conducting a game with >2 players. It very much depends on the size of your lab. The idea here is something like a design of subtraction: we want to prevent participants from thinking that they are playing repeatedly with the same individual(s) although anonymity in the strangers design mitigates it somewhat. Longer term relationships differ from short term ones and may affect behaviour via several mechanisms.

For the stranger matching, a key parameter here is the number of rounds relative to the number of participants in the lab. If the number of participants were very small, then conducting many periods with stranger implicitly means that they will be playing with others very often. One needs to keep this in mind..

5. **Deception offers many advantages in the laboratory so why is it not typically used by economists?**

Having to tell the truth to participants is a constraint on the experimental design, so it makes examining certain issues more difficult. See the psychology literature (e.g. Milgram experiments) for examples of how deception can be used.

Problems with deception

Using deception in the laboratory is problematic when subjects eventually realise that they have been deceived (Note that deception experiments require a post-experiment briefing). In future experiments subjects will then be less likely to believe what they are told (assuming the pool of subjects remains partly static) and this introduces problems when analysing results. Having a norm of no deception and being clear on this to subjects is therefore a useful norm within experimental economics.

Note that the participants which we elicit in the lab are relatively well experienced in experiments (This is a much bigger problem in Amazon Mechanical Turk), and this already constitutes a problem since they may be guessing what the experimenters want of them (experimenter demand effect). If in addition, they start to second guess your experimental procedure, this will create more noise.

For example, in interactive experiments, they may be wondering whether they are playing against real participants, conspirators, or bots. *This will result in differences in the way they interact to the extent that perceived beliefs about others matter (guilt, disappointment, betrayal aversion, social norms etc.) Betraying a computer has no moral implications for example..*

If you use the Economics lab, it might be useful to emphasise that no deception is used for our experiments; in emails sent out they do mention this. In my experience, some people have been second guessing experiments which I have conducted: whether they are chatting with real people, or whether I am truly assigning them based on what I have said.

Deception vs Omission

On the other hand, Economics does allow for hiding the full truth from participants. For example, in multi period games, if we do not want there to be end game effects, we can omit the number of periods we conduct and just say: "You will play several periods of the game". In the stranger matching as before, they might play with the same individual twice, but we do not explicitly say this: but only say: "You will be randomly matched with someone in the session". In contrast, when using perfect strangers matching: we often say: "You will be randomly matched with someone in the session AND will not play with the same person twice". Basically, the truth is told, but possible harmful inferences are omitted in instructions.