Behavioral and Experimental Economics

Lab Experimental Economics in action: Exam rules and simulation

Paolo Crosetto

The role of behavioral public policy

Getting into the mind of subjects

- focus on cognitive aspects
- clearly identify mechanisms
- (if needed) sidestep preferences
- heuristics, choice processes

Building counterfactuals

- explore different scenarios
- integrate preferences with control
- track macro consequences
- cheaply explore solutions

Your exam– getting back and forth

Take-home 4-pages report a paper

The article report is made up of **three parts**:

1. 1-page summary of the paper and main results;
2. 1-page critique of experiment: what are the weak points? Does it lack in external/internal validity?
3. 2-page alternative design proposal: transition from *mechanism* to *counterfactual*
   - if the paper studies a *mechanism*, propose a **counterfactual** study (lab or field)
   - if the paper studies a *counterfactual*, propose a **mechanism** study (lab or field)
   - the design includes a data analysis plan, that roughly describes what data and what analyses you'll carry out on the paper.

More details in the Exam Rules pdf on github

Your exam: deadlines and practical stuff

1. The exam is a 4-pages take-home *paper report*.
2. Papers come from the reading list discussed on April 28th
3. **Due on June 15th, 2023**.
4. Worth **30%** of the final overall mark (Anna's part is worth 70%)

Assigning papers

Assignment is managed by yourselves on this google spreadhseet.

Report structure

- Summary of the article: 1 page
- Critique of the design: 1 page
- Proposal of a new design including data analysis plan: 2 pages

Let’s look at the papers

How to write the report?

Part 1: Summary

This one is easy. But don't forget to include:

- What is the research question?
- Which typo of experiment is this? (if it is one)
- What methods are used?
- Details about the sample
- Details about results
- Limitations of results

Different types of experiments

You can hence choose your pick of experiment

Natural: the exogenous manipulation and the control group happen randomly and naturally in the population. The expeirmenter can simply look at the real world data [very rare!]

Field: the experimenter creates a control group and an exogenous manipulation, randomizing some aspects and keeping others as they appear in nature. [also called Randomized Control Trial]

Lab: the experimenter takes full control and *recreates the setting in a lab*; synthetically creates a control group and a manipulation, and sees details [this is the gold standard in the hard sciences]

Natural experiments

If you are lucky enough, nature provides you with a ready-made experiment

Conditions:

- the manipulation must be *exogenous* (i.e. out of the control of the agent)
- there must be no selection (i.e. subjects being able to opt-in or -out)
- subjects must be unaware of the manipulation
- the manipulation must fall across lines that while not random, are so from the point of view of the agents

Example (Crime rate decline in the US and abortion laws)

According to Steve Levitt (book: Freakonomics) the legalization of abortion in the US reduced crime 20 years later. How can he prove it? via a natural experiment. Some US states legalized abortion before the others – and saw earlier declines. Moreover some states made it harder than others to perform an abortion – and in those states crime declined less. [more details on the video]

Field experiments: Randomized Control Trials (RCT)

If you are a medical doctor, what you do is RCT

- Take a population fo interest (e.g.: lung cancer patients)
- administer a drug to a treatment group, a placebo to a control group
- track effects of the drug
- done.

Control vs. Randomization

- The setting, drug, background health levels, conditions of exposure to the drug, etc... are controlled
- The sample, number of patients, hospitals, control/treatment, are all randomized

But can it be generalized to other patients? other settings? slightly different drug? not really..

Lab experiments

High internal validity, low external validity

- Recreate in a lab the *essential traits* of the setting to be studied
- That is: simplify, simplify, simplify!
- Keep only the minimal things that can prove your theory wrong
- Formally create one (or more) control groups
- done.

Control vs. Randomization

- All is controlled, synthetic, recreated in perfect (from the point of view of the theory) conditions
- Only allocationt o treatment or control is randomized

Part 2: critique

This one is trickier. Look in detail at slides from our first session.

- Internal validity?
- External validity?
- Possible confounds?
- Causality
- Sample and selection effects
- How much confidence do you give to results?

Internal validity

Definition (Internal Validity)

In scientific research, internal validity is the extent to which a causal conclusion based on a study is warranted[source: Wikipedia]

Example (Impact of advertisement)

A national advertisement campaign runs on TVs from 10 to 15 february, 2017, to incentivise use of buses. In the second part of the month you observe an increase in the use of buses. Did the campaign work?

- temporality: purported cause should come *before* the effect
- correlation: is there a correlation between campaign and outcome [yes]
- causality: need to check
  - if the result really exists [in line with 2nd half march 2016? 2015?]
  - if other factors could have led to the result [good/bad weather? discounts? pollution peak?]
  - if some non-controlled group entered your observation [big meeting of hi-school students in Grenoble]
  - if the very fact that you measured has increased attendance
  - if a group had some interest in the campaign appearing to succeed and act accordingly
- if *one* of the ifs above is not fulfilled *then* you do not have internal validity

Possible confounds: confusion

Definition (Confusion effects)

Biases introduced in your experiment by subject not having understood the experiment and interacting with other subjects, hence biasing the whole session

Example (Trading permits in the EEDD master class)

When we traded permits, some of you bought a permit at a way too high price – because they had misunderstood the rules of the game. Consequences go *above and beyond* simple noise (for which statistical techniques exist to take care of)

- Subjects bid too high and as a consequences win
- Still, they make losses
- By bidding too high, they crowd out of the market other bidders
- In the end, all the trading fails, permits are not used, no units are exchanged
- The misunderstanding creates a cascades of consequences also in the behaviour of *other* subjects

Possible confounds: experimenter demand effect

Definition (Experimenter demand effect)

Biases introduced in your experiment by the fact that the subjects act as if to do what they believe the experimenter wants of them, 'the right thing', 'the good thing'.

Example (Energy labeling)

You run an experiment comparing different energy labels on washing machines. You let the subject understand that out of the different labels, one was designed by the experimenter himself. Changing electricity consumption is not so difficult or costly (e.g.: unincentivized questionnaire). Subjects might deflate their declared consumption for *your* label because they want to *help* you.

- are we observing a real behavior or an artifact?
- if we prove causality, can we trust the proof?
- ways around:
  - neutral instructions
  - context independent choices: anonymized objects and labels
  - blind testing: the experimenter is a third person

Possible confounds: incentive size effects (including size = 0)

Definition (Incentive size effects)

Biases introduced in your experiment by the fact that the small (or null) monetary values present in the experiment elicit qualitatively different behavior than the real-world high stakes

Example (Giving in trust games)

You run a trust game, in which a subject has an allocation of 3 euro, has to split it with another subject; any amount split is tripled, and the receiver has the ability to share it back (or not). Equilibrium is for the first subject to send zero; you observe people sending the full three euros.

- would this happen with complete strangers in the road and 300 rather than 3 euros?
- is the result an artifact of the very small incentives at stake?
- (note that things get worse for questionnaires: it is *free* to lie / misreport / make mistakes)
- Possible solutions: give appropriate incentives; prove behavior is unchanged as incentives change.

Possible confounds: sample selection effects

Definition (Selection bias)

Selection bias is the selection of individuals, groups or data for analysis in such a way that proper randomization is not achieved, thereby ensuring that the sample obtained is not representative of the population intended to be analyzed [source: Wikipedia]

Example (An unexpected sample)

You expose 50 consumers to a questionnaire on their use of electricty. Turns out the most used means of transport is the bike, and 30% have photo-voltaic cells on their roofs. You meant the sample to be random, but due to word of mouth 78% of subjects come from a group of activists of the green party.

- selection effects can lead to severely biased and false claims (made in good faith!)
- selection is everywhere: M2 class is NOT a random sample of the student population
- if you have selected subjects, you only observe *conditional* probabilities
- and these cnanot be generalized
- solution: randomize, check, randomize again, use large numbers, pay attention!

Possible confounds: lack of credibility

Definition (Experimenter credibility)

The fact that experimental subjects believe the instructions to be the real rule of the game. Control about the *beliefs* the subjects have of what is going on in the laboratory.

Example (Electrocution)

A famous Social Psychology experiment involves questioning a subject telling him that he can choose whether to earn some money or not, knowing that if he does then in an adjacent room another subject will be (lightly) electrocuted. The experiment investigated the moral sense of people faced with immoral rules of the game. Nobody was really electrocuted, but a confederate (an actor acting on behalf of the experimenter) was making pain noises from an adjacent room.

- experimental results are strong only insofar as the scenario is believed
- that is, the script must be fully credible
- taking the money does not only mean not caring about the other suffering; it might also mean that you do not believe anyone is really suffering.
- credibility is key: beliefs guide actions and all we can observe are actions (and not beliefs!)

External validity

Definition (External Validity)

In scientific research, external validity is the degree to which it is warranted to generalize results to other contexts. [source: Wikipedia]

Example (Impact of advertisement)

You expose 45 subjects to a campaign to incentivise the use of buses.They report to you in a questionnaire that after seeing the ad they feel more inclined to take the bus by an average 20% more. Will the ad work also in the field?

- representativeness of the sample [to represent different types of people]
- size of the sample [to exploit statistical inference]
- control group [to ascertain effect different from baseline or random]
- incentive compatibility [to make sure that subject do not lie to you or to themselves]
- double-blind testing [both the experimenter and the subject are blind as to the purpose of the test]
- more...
- if *one* of the ifs above is not fulfilled *then* you can arguably have *low* external validity

Part 3: new design & data analysis

This one is the hardest one. But you get some guidance!

- think outside of the box
- mechanism to counterfactual:
  - move from lab to field
  - or from lab-mechanism to lab-consequences
  - think about impact on the market at large
  - think about real-world interventions
  - keep control!
- counterfactual to mechanism:
  - move from field to lab
  - google is your friend: look for potential candidates to explain the field
  - usually field papers apply lab papers: go look for the source
  - ask yourself "why this happens"?
  - go abstract, obsess about control, reality be damned

How to get external validity: get near to theory

How do you get external validity?

Get as far away from reality as possible and near to theory

- theory is generalizable by default
- so the closer your experiment is to theory, the higher the chances of falsifying it
- if you fail, theory survives and you can apply it to further contextualized settings
- because imagine theory fails in *one* little setting
- then it can be because of some confounds, some external influence, while theory is still valid
- if the test is non-contextual then it is in principle stronger

How to get external validity: get near to the real world & replicate

How do you get external validity?

Get as much real context as possible, and then replicate in several different contexts

- running an experiment in the field, within the target population, fully contextualized
- makes sure the result is valid in *that specific field*
- *but* leaves open the possibility that it will *not* replicate in *other* contexts
- from which we deduce the need for replication

An example of different methods: Electric appliances

Energy efficiency gap: consumers do not buy the best energy-efficient appliances even if they (and the environment) would benefit Which interventions work?

- Several RCTs failed to give a definite answer on what works and what not.
- See the paper by Pete Lunn on Moodle

Country	Appliance	Intervention	Incentives?	Works	doesn't
Germany	Washign machines	info on lifetime operating cost	yes	X	
Ireland	Tumble dryers	info on 5-year energy cost	yes		X
Finland	Tumble dryers	energy label + staff training	yes	X	
Finland	Tumble dryers	energy label only	yes		X
Finland	Tumble dryers	staff training only	yes		X
Finland	Fridge freezers	energy label + staff training	yes		X
Germany	Televisions	info on lifetime operating cost	no	X	
Germany	Televisions	info on yearly operating cost	no		X
UK	Boilers	info on yearly operating cost	no		X

Hard to find a pattern – each RCT is valid but *limited* in scope

- Would lab experiments be better? Yes, if successful in finding the underlying mechanism
- e.g.: problems in processing numbers; ratio of upfront to operating cost... if the underlying mechanism is understood, policy follows

Data analysis plan

This is tricky but doable. Choose among:

- between subjects – gold standard, but high N and no before-after
- within subjects – order effects, but before-after
- diff-in-diff
- complex structures

And think hard about:

- moderators
- control variables
- subgroups
- which tests to run

Let's do this together: antisocial punishment

The public good game

Let us generalize this to N players: public good game

Rules of the game

- Each of you has a (fictitious) endowment of 20 euro
- You play in groups of **4** players (fixed matching)
- each of you has two accounts
  1. a **private** account, that returns 1 euro for each euro invested (by yourself only)
  2. a **public** acccount, whereby each euro invested there is multiplied by 1.6 and then shared equally with all other players
- The payoff in each period is simply the sum of the earnings from the two accounts
- your decision is **how much to contribute to the public account**

play now https://classex.uni-passau.de/ class: INRAE code: M2EEDD

The public good game

Public Good Game: mechanism

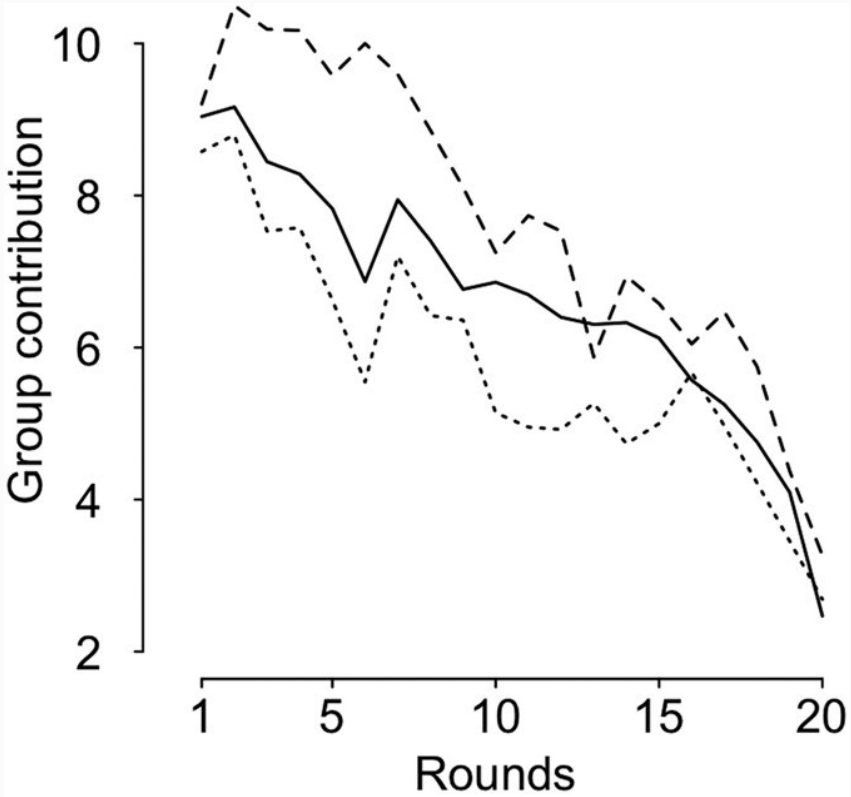
- Mechanism is exactly the same as in Prisoner Dilemma
- There is an action that generates public benefits...
- ...but at a private cost!
- the *social optimum* is given by everyone contributing everything
- yet, *individually* it is a dominant strategy to contribute *less* than the others, for any level of the other players' contribution

The only Nash is for everyone to contribute zero

The public good game

Public Good Game: results

- contributions usually start off quite substantially above 0
- but then decay with repetitions, usually ending at around 0



[Tognetti et al: http://www.nature.com/articles/srep29819]

The decay in contributions

Why this decay in contributions?

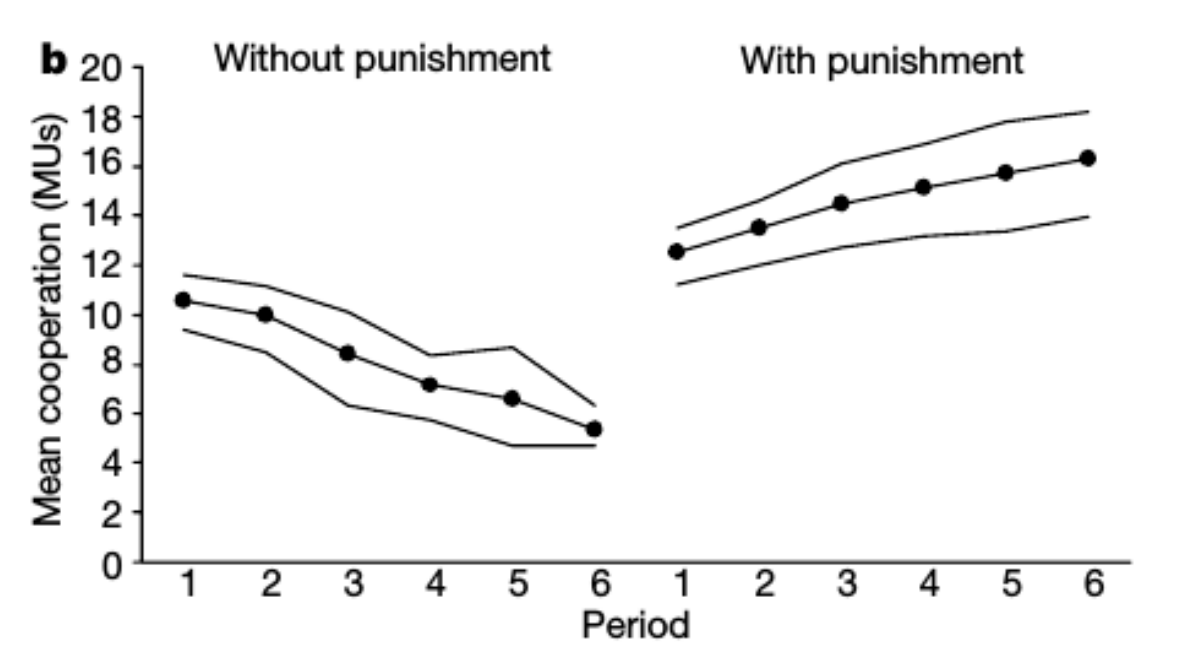
- some people are free riders, some cooperators
- but most people are *conditional* cooperators
- they are happy to contribute, but they do not like being cheated by the others
- if others lower their contributions, they do too
- leading to cascades of negative reinforcement
- and finally to very low contributions

This is the main reason why voluntaristic endeavours are often unstable and short-lived

Punishment

- In society, police exists to enforce rules
- it is costly: we have to pay for it
- would it be possible ot have endogenous punishment – i.e., to have no police but to rely on peers to sanction each other?
- Altruistic punishment: each subject has a right to sanction others
- but this is costly: subjects *pay a fee* to sanction others – i.e. reduce their payoff
- for instance, you can burn another player's money at a cost of 1/3 of euro per each euro burned
- it is *irrational* to do so – it costs you money!

Punishment: results



[Fehr and Gächter, Nature 2002]

Is punishment effective?

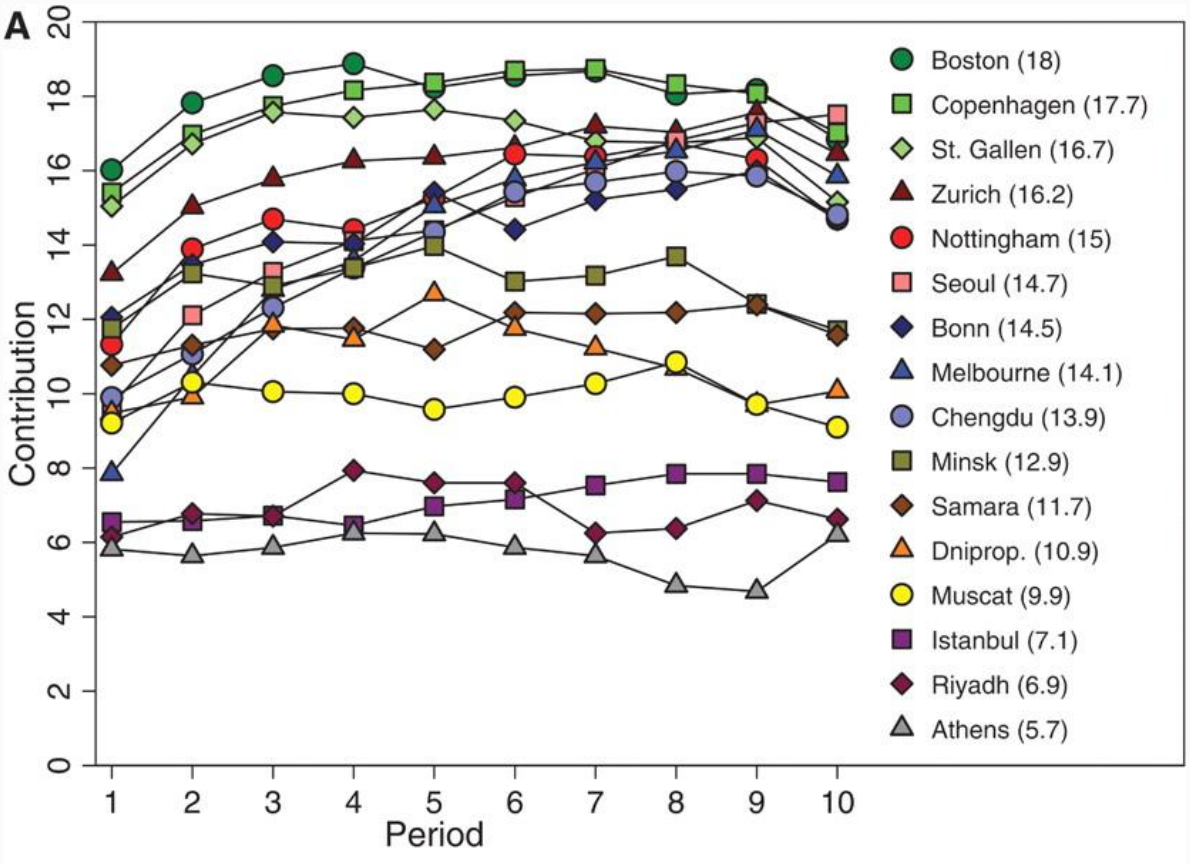
Yes:

- rule self-enforcement works very well
- sometimes without the need of actual enforcement: the threat suffices
- thus only mildly affecting welfare

But:

- it depends on *willingness to enforce* on the part of subjects
- ...and it depends on *which rule* subjects want to enforce
- if the rule is anti-social in itself, *that* will be enforced

Punishment not working



Antisocial punishment

What if subjects punish the *good guys*?