INTRODUCTION

Why Experimental Economics?

- In empirical research, natural sciences rely almost exclusively on (more or less) controlled experiments. Examples:
  - Large Hadron Collider in particle physics
  - Laboratory and clinical studies in biology and medicine
- Historically, economics (like astronomy, for example) has relied on observation of naturally occurring processes for its data needs.
  - Advantage: these processes are observed in their undisturbed form
  - Disadvantage: it is difficult to establish causal relationships among observed variables since many of them are determined simultaneously (example: wages of married men)
- In late 1940s (Chamberlin) and in 1950s (Smith), this motivated researchers to start conducting economic experiments in the lab.
- Advantages of economic experiments in the lab:
  - controlled environment, it is clear what the cause and what the effect is;
  - ability to observe counterfactual scenarios;
  - ability to test novel designs of economic institutions such as markets.
- Disadvantages of economic experiments:
  - problems with **external validity** (the degree to which results obtained in the lab approximate what happens or would happen outside the lab);
  - non-representativeness of subject pools;
  - heterogeneity in individual preferences and cultural or social norms across space and time;
  - potential unobserved loss of control (e.g., experimenter effect, subject pool contamination, weakness of incentives, etc.)
- A recent development that tries to address some of the disadvantages: field experiments
  - improved, although not perfect, external validity (depends on whether subjects know they are participating in an experiment);
  - subject pools likely to be more representative, although depends on the design.

Why do we run economic experiments?

1. Testing theory/establishing causality
    - this is the original motivation behind running controlled experiments
2. Fact-finding
    - for example, investigating gender differences in attitudes toward risk and competition
    - for example, comparing efficiency of various market institutions
3. Whispering to the ears of princes
    - this involves work that is supposed to have direct policy implications
    - for example, it could involve design of auctions for mobile spectrum licenses

HISTORY OF EXPERIMENTAL ECONOMICS

- The discipline started with the work of Richard Chamberlin in late 1940s and Vernon Smith in 1950s
- Very few articles published until late 1970s (less than 25 a year)
- Significant growth in publications to above 100 a year by the end of 1980s
- More than doubling to above 200 a year by the end of 1990s
- Growing recognition in the profession:
    - Nobel prize awarded to Vernon Smith (experimental economist) and Daniel Kahneman (experimental psychologist) in 2002
    - Nobel prize awarder to Reinhard Selten (game theorist and experimental economist) and John Nash (game theorist) in 1994, although mostly for their contributions to the development of game theory

METHODOLOGY

Basics

- We are usually interested in a causal impact of an exogenously controlled set of procedures, instructions, incentives, rules and parameter values, called a **treatment**, on an endogenous variable capturing some aspect of subject behavior, called an **outcome**.
- Examples:
    - Treatment: size of the pie to be split in a dictator or an ultimatum game; Outcome: amount retained by the person making the offer.

- Treatment: origin of the pie to be split in a dictator or an ultimatum game (manna from heaven vs. earned); Outcome: amount retained by the person making the offer, rejection probability.
- Treatment: auction format (first-price, second-price, English, Dutch); Outcome: auction revenue, auction efficiency

- We are sometimes interested in treatment effect of more than one aspect of the environment. For example, in the ultimatum game it can be the effect of the size as well as the origin of the pie. In that case, researchers usually implement what is known as a **factorial design**: if m different treatments are considered for the first and n for the second aspect of the environment, then run treatments in all $m \times n$ design cells. Factorial design may also involve more than two dimensions:
  - Example: (first-price vs. second-price auction) x (with known vs. unknown distribution of valuations) x (with or without reserve prices)

- This causal impact is usually referred to as the **treatment effect**. In an ideal situation, treatment effects are identified by systematically varying the relevant treatment, each time observing and recording the corresponding outcome, holding all other relevant factors/variables constant (*ceteris paribus*). That way any observed change in the outcome can be attributed to corresponding changes in the treatment, and hence causality can be established. Even though we never live in an ideal world, the first and most important rule of experimental design is: **change only the treatment, holding all other potentially relevant factors constant.**

- What are all such other potentially relevant factors that need to be kept constant?

  1. other features of experimental design and implementation (including the physical location of the experiment for different treatments)
  2. experimenter and his/her attitude
  3. subjects and their mindset

Order Effects, Within- and Between-Subjects Designs

- It is relatively easy to do (1) and (2), but (3) is much more difficult. The reason is that if the same set of subjects is sequentially presented with alternative treatments, their behavior is going to be affected by so-called **order effects**, meaning that they may behave differently in the same treatment depending on where in the order it comes. Order effects may be present due to, among other things,

  1. experience from previous treatments
  2. subject fatigue
  3. income effect from earnings accumulated in previous treatments

- There are two usual ways of dealing with the first two sources of order effects:

  1. Use different subjects for different treatments, where subjects are **randomly assigned** to different treatments. At the first inspection, this seems to go against the spirit of "same subjects with same mindset." This approach is, however, based on the assumption that any randomly chosen set of subjects will have, on average, the same personal characteristics background. As a result, there will be no

correlation between these personal characteristics and a particular treatment and one therefore can, with sufficiently many subjects (Law of Large Numbers) obtain a relatively precise measure of the treatment effect. Use of different subjects in different treatments is referred to as **between-subjects design**.

2. Use the same subjects for different treatments, but randomize the ordering of treatments across different experimental sessions. Use of the same subjects in different treatments is referred to as **within-subjects design**.

- An advantage of the between-subjects design is that is does indeed eliminate order effects in a clean way. Compared to that, randomizing ordering of treatments in the within-subject design is a heavy-handed way of dealing with the problem as it simply attempts to average the order effects out rather than to eliminate them. On the other hand, the within-subjects design gives the best possible control for subject personal characteristics across different treatments, equivalent to the possibility of using **fixed effects** in regression analysis. Compared to that, between subjects design achieves control for these personal characteristics only by averaging and relying on the Law of Large Numbers. Hence it often requires a higher number of observations for the same precision (standard error) of estimates of the treatment effect. This is especially true if variation of such personal characteristics within the subject pool causes a large variation in the outcome that is not related to variation in the treatment.

  - Example: running speed in shorts vs. jeans when subject pool includes a representative subset of the population.

- An important issue when running repeated sessions and hence almost always in between subjects design is that **no subject participates more than once**. This is important from the statistical point of view since in order to be able to invoke the usual Law of Large Number properties, you need independent observations at the level of subjects and/or sessions.

- In order to deal with the third possible source of order effects, if possible, **payoffs from tasks should not be revealed until the end of the experiment** so that behavior is not affected by previously realized earnings. A popular choice to achieve this is called a **strategy method** under which subjects specify their behavior in various possible scenarios of the moves by nature and other players (think of an equivalent of a complete contingent plan from game theory), and then a subset of these scenarios are implemented and subjects are paid based on payoffs realized in these scenarios. **Holt-Laury risk aversion elicitation tool** is an example of use of the strategy method. This method is also popular when an experimenter wants to implement high stakes, but, due to budgetary considerations, can only pay subjects is some rounds or can only pay some subjects (above the participation fee).

Importance of Monetary Incentives

- It is a widely agreed practice in experimental economics that subjects should be incentivized by real (as opposed to hypothetical) money. This is because real-world economic decisions usually involve money or utility differences, and utility differences can in turn be converted to money differences (compensating and equivalent variation).

- There is some debate about two aspects of paying subjects:

  - cash vs. non-cash (such as bank transfer)

- immediate post-experimental payment vs. delayed payment vs. pre-payment of some portion of the reward

- Due to salience reasons, the most usual way subjects are being paid is **in cash at the end of the experiment**.

Other Important Design Issues

- **Do not use deception!** The reason is that the lab needs to have a reputation for being faithful to the instructions and not deceiving. This is because otherwise subjects, in expectation of deception, may behave in ways that they would not behave if they fully believed the experimenter. That is, a reputation of the lab for not deceiving is an important public good, and each experimenter who uses the lab should contribute to it.

- Tips on instructions:

  - Draft instructions that are detailed enough for subjects to understand but not overwhelming to grasp.
  - (question of taste) Provide subjects with a printed version of the instructions for their reference anytime during the experiment.
  - Draft the instructions in such a way that you can follow them *verbatim*, without any need to provide unrecorded clarifications or examples. This is important for maintaining control across different sessions and for **replicability**. One aspect of this is to minimize the number of idiosyncratic questions asked by subjects, since such questions may affect behavior of other participants in the session and hence entail a (partial) loss of control. Particular advises are:

    1. Include **examples** and/or **practice questions** and/or **practice rounds** that give subjects a hands-on way to practice the task before the paid rounds begin. But design these carefully such that they do not suggest behavior to subjects, i.e., avoid **anchoring effects**.
    2. Use a **neutral language** and **avoid using emotionally loaded terms** (unless purposefully part of the design) and economic jargon (if it has not been previously defined/explained). Examples:

    * · Instead of "...another subject may steal your ...", use "...another subject may take away your..."
      · Instead of "...may allocate tokens unfairly..." use "...may allocate different amounts of tokens..."
      · Instead of "... the marginal benefit of the additional unit..." use "... an additional benefit of using one extra unit..."

    3. Provide a useful amount of **context** (using words such as market, demand, supply, bid, reserve price, payment), but avoid contextual labels that are not important or may distort the results.

- (question of taste and need) Follow the experiment with a **questionnaire** in which you collect valuable demographic information such as gender, age, academic major, and whatever else you think is valuable for your project.

- Record a **lablog** during each session that you run. This may later help you to understand some quirks in the data.

- Carefully document all aspects of experimental design and implementation. Do this because:
    1. these records will be useful for you later on when you analyze the data
    2. journal editors often ask you to make experimental details and data publicly available as a condition of publication
    3. this is a good practice for the purpose of **replicability**

Logistical Issues

- Plan ahead: it takes time to design an experiment, program it, test the program, run a pilot, recruit subjects and run sessions.
- Before running a pilot, **dry-run the programs** in the lab, trying to input crazy things, to make sure everything works as intended.
- Before running experimental sessions for real, run at least one **pilot session**, preferably using your colleagues/classmates/students as subjects, and ask them for an extensive feedback on the design. Incorporate reasonable suggestions.
- Try to minimize **subject pool contamination**. This refers to a situation when some subjects that are just about to participate in your experimental session already have an idea of what will be going on because they heard about it from earlier participants or other sources. Here are some of the ways to try to minimize subject pool contamination:
    1. Use a separate subject pool for the pilot session(s).
    2. Try to run the real sessions in as short a time span as possible.
    3. If different sessions involve different degrees of information revelation to subjects, run the lower-information sessions first and higher-information sessions second.
- Use an online recruitment tool such as **ORSEE** to recruit your subjects. This makes things simple. You need to **over-recruit** somewhat to make sure your sessions are not ruined by **no-shows**.
- In the recruitment ad, specify the **show-up fee** (paid to subjects who show up but do not end up participating). This may be in the range of 100-150 CZK. You also may, but do not have to (depends on lab rules) specify a **minimum payoff** from the experiment. This may be potentially problematic because it gives underperforming subjects increased incentives to take risk.
- Get your cash ready in **sufficiently small change**.
- Although this may seem excessive, a usual problem when running multiple sessions is that some subjects try to participate more that once, which disturbs statistical independence of individual subject and/or session observations (see above). As a result, I would recommend that you ask the subjects to present a **picture ID** before the experiment.
- Before the experiment, you should have subjects sign a **Subject Consent Form** that very briefly outlines the experiment, in which they affirm their agreement with participation in the experiment. This is to protect yourself against possible later complaints by subjects that they were misled in some way.
- When you pay the subjects after the experiment, have them sign a **receipt**. You will need to present this to the Accounting Department in your institution.

Last Message Before You Go and Design Your Own Experiment

- Designing and implementing a good experiment is to some extent an art. So good luck!