Pratical matters

- My name is Guillaume Hollard. I am a researcher at CNRS and Ecole Polytechnique.
- Feel free to contact me for questions at guillaume.hollard@polytechnique.edu
- We have a total of 24h, (imperfectly) divided among 9 or 10 classes depending on the evaluation
- Evaluation will consist in an oral presentation and/or a written examen and/or a written essay and/or running your own experiment. The evaluation format strongly depends on the number and motivation of students
- Once the registration time is over (by early November), we will jointly choose the binding format
Participation!

- I strongly encourage participation!
- We will play games, comment on some results, try to have a critical approach, etc
- All theses aspects require active participation.
- Furthermore, there are some topics to be chosen (cf syllabus)
- Last, for each chapter I’ll try to indicate the main questions raised. I expect you to be able to answer such questions
Introduction

- Experimental economics refers to the use of experiments in economics
- The first use of experiment in economic is rather old (several attempts in the 30’s)
- However, much of the action happened in the recent years with an incredible boom of research in this area
- As a result, there is now a continuous flow of new papers that have transformed the field a lot
- The purpose of this class is thus to make sense of this rapidly growing literature by providing a big picture of what is going on
The big questions

- What is experimental economics?
- What is experimental economics good for?
- What are the limits?
- What are the hot topics?
- Where are we going to?
Today’s questions

Today, I would like you to pay attention to the following things:

- The main topics in experimental economics (it would be good to have at least a rough idea of what each field is about)
- The diversity of experiments (methods, subjects pools, goals, etc)
- The role of experiments in science
In one classic 1882 example, Louis Pasteur designated half of a group of 50 sheep as controls and vaccinated the other half. All animals then received a lethal dose of anthrax. Two days after inoculation, every one of the 25 control sheep were dead whereas the 25 vaccinated sheep were alive and well!

- In that example, a simple and convincing experiment seems enough to prove causality (it is the vaccine that creates immunity)
- Transferring the observed effect from sheep to humans may seem a good idea (and it indeed was)
Experimental economics: the core idea

- The idea at the very heart of experimental methods is to randomly assign subjects to groups (e.g. creating two groups of 25 sheep each)
- The experimenter has control on the variables that creates difference among groups (e.g. the experimenter vaccinates sheep in one group)
- The differences in outcomes (e.g. dead or live) can thus be attributed to the variables that are under the control of the experimenter
- Experiments allow the researchers to generate the counterfactuals (i.e. what would have happened if...) we are (sometimes desperately) looking for in economics and social sciences (think of any political debate...)
Experimental economics: the core idea

The basic idea is very simple and intuitive and there is not much more behind experimental economics!

Some questions:

▶ Why is randomly creating the groups so important? (cf selection bias)

▶ To what extent does this amounts to creating several copies of the world? (cf hidden variables, size of the sample, etc)
Is running experiment simple?

- The core idea is nice and simple but...
- Treatment effects are not always simple to identify, e.g. the effect may have different effect on statistically similar individuals
- The nice thing with the sheep example is that all sheep in the same treatment reacted the same way!
- Even when all precautions are taken, there is not always one single interpretation of the difference across treatments (e.g. which component of the vaccine creates immunity)
- Some experiments are more convincing than others...
Our world is a messy one...

- As a consequence, there is not one best way to design a good experiment, i.e. what is a good experiment is not straightforward.
- There are clear mistakes not to be made, but nothing like a set of rules that will ensure that the experiment will be a success.
- For those who do theory, I guess it is a fair account to say that the main difficulty is to struggle with mathematical problems so as to prove that the expected result holds under general enough conditions.
- In empirical work (not only experimental economics) the trouble comes from the fact that the data rarely deliver a simple, non-ambiguous message.
- Getting more control (e.g. lab experiments) allows us to remove some of the "noise"...
- But the price to pay is that we study artificial situations...
Social and "asocial" sciences: a comparison

- Pasteur: All animals then received a lethal dose of anthrax. Two days after inoculation, every one of the 25 control sheep were dead whereas the 25 vaccinated sheep were alive and well!
- Typical results in economics: Two sheep refused to receive anthrax, even for a reasonable amount of money. One sheep did not understand the purpose of the experiment and was excluded from the analysis.
- Most sheep are studying economics as a major.
- 75% of vaccinated sheep did survive, 42% of non-vaccinated sheep did survive. The difference is statistically significant at the 5% level, so we can conclude that the vaccine is effective.
Let’s have a look!

- Rather than long talks about experiments, I suggest we rather review some experiments together.
- We will look at three experiments.
- I will not try to provide many details, just the general picture.
- However, if you feel some relevant information is missing, feel free to ask.
- I choose these experiments because they differ on many aspects (methods, subjects, purpose, etc).
Looking for facts: The marshmallow experiment

- This experiment consists in presenting kids with a reward (the marshmallow)
- and to offer them a greater reward if they were able to wait long enough (say 15 minutes).
- Here is a video that nicely summarized it: http://www.youtube.com/watch?v=4ZikfUI0G5o
- We will watch it if the gods of technic allow us
The marshmallow experiment: the original Mischel experiment

- The original experiment took place at a Nursery School located at Stanford University in 1970.
- It uses children aged four to six as subjects.
- Children were presented with a reward (marshmallow, cookie, or pretzels as in Mischel).
- The children could eat the reward, the researchers said, but if they waited for fifteen minutes without giving in to the temptation, they would be rewarded with a second, identical, reward.
- In over 600 children who took part in the experiment, a minority ate the marshmallow immediately. Of those who attempted to delay, one third deferred gratification long enough to get the second marshmallow.
The marshmallow experiment: follow-up studies

- Since Mischel’s daughters knew and grew up with many of the original test subjects, through casual conversation, Mischel discovered there existed an unexpected correlation between the results of the marshmallow test, and the success of the children many years later.

- In particular, the ability to delay gratification also correlated with higher SAT scores. The one marshmallow kids scored an average of 210 points less on SAT tests (it has thus a major impact about 15 years after)

- A 2011 brain imaging study of a sample from the original Stanford participants when they reached mid-life showed key differences between those with high delay times and those with low delay times in two areas: the prefrontal cortex and the ventral striatum when they were trying to control their responses to alluring temptations.
Defriefing: marshmallow

- What do you think about that experiment?
- In your view, it is an interesting experiment?
- What have you learned?
- What kind of follow-up study would you like to carry on?
- Is it important for economics (i.e. should we care?)
Economics provides a lot of models aiming at predicting "rational" behavior in various circumstances (games, decisions, ).

In particular, we can test whether the axioms and hypothesis used in economic models are satisfied.

You probably heard of some of these experiments (e.g. the Allais paradox, the ultimatum game, ).

I will here present some of these experiments.

If you already know these please let those who don’t enjoy them!
Let’s play

- In what follows, I will present several decision problems
- I will ask you to make some choices
- It would be nice if you have a small piece of paper ready
Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Rank the descriptions below in term of likelihood they resemble Linda.

- a. Linda is a teacher in an elementary school.
- b. Linda works in a bookstore and takes yoga classes.
- c. Linda is active in a feminist movement.
- d. Linda is a psychiatric social worker.
- e. Linda is a member of the League of Women Voters.
- f. Linda is a bank teller.
- g. Linda is an insurance salesperson.
- h. Linda is a bank teller who is active in a feminist movement.
Bob witnessed an accident: he saw a blue cab creating an accident. The cab hit and run at night. The court tested whether Bob was indeed able to distinguish blue cabs from green cabs at night (in that city 85% of the cabs are Green 15% of the cabs are Blue). They estimate that Bob is right 80% of the time (I am fine with an immediate answer, no computation required thus). What is the probability that the cab is Blue?
Results: Bob

- Majority says 80% because they think witness’ judgment is representative of the color of the cab: \( P(\text{identify B/ B}) = 0.8 \)
- Subjects ignore the base rate 0.15 of the cabs being blue
- Accounting for the base rate gives:

\[
P(B / \text{identify B}) = \frac{P(\text{identify; B/ B})P(B)}{P(\text{identify; B/ B})P(B) + P(\text{identify; G/ G})P(G)} = 0.41
\]
A group of 88 undergraduate students at UBC ranked the 8 statements. The description of Linda was constructed to be representative of an active feminist (F) and unrepresentative of a bank teller (T). So in fact only three of the statements are useful for that experiment

- c. Linda is active in a feminist movement. (F)
- f. Linda is a bank teller. (T)
- h. Linda is a bank teller who is active in a feminist movement. (T&F)

85% of the subjects indicates that $P(F) > P(T&F) > P(T)$

What do you think about it?
Linda: results

- A very low probability of Linda being a bank teller, $P(\text{Linda is a bank teller}) = 0.05$
- A high probability that she would be a feminist, $P(\text{Linda is a feminist}) = 0.95$
- Assuming independence: $P(\text{Linda is T and F}) = 0.05 \times 0.95 = 0.0475 < P(\text{Linda is a bank teller})$.
- Why? A detailed, specific scenario seemed more likely because of the representativeness heuristic, but each added detail would actually make the scenario less and less likely
Defriefing: Linda and Bob

- What do think about these experiments?
- In your view, are these interesting experiments?
- What have you learned?
- What kind of follow-up study would you like to carry on?
- Is it important for economics (i.e. should we care?)
Matching markets are specific markets in which it is not only about choosing (like in a supermarket) but also about being chosen.

There are countless such markets, including the job market for new entrants but also marriage, kidney donation, etc.

A prominent figure in this area is Al Roth who received the Nobel Prize in 2012 for his work on matching.

In what follows, we will describe and analyze how he uses experimental economics to study a particular phenomenon called "unraveling"
Let us consider two market institutions:

- A centralized market, associated to a "clearinghouse" in which offers from both sides are presented and treated simultaneously
- A decentralized market with no constraint on the timing of contracting

Many (decentralized) markets have trouble coordinating on the timing of transactions. There are many empirical examples of unraveling markets characterized by:

- Offers which become progressively earlier as participants seek to make strategic use of the timing of transactions
- Generalization of exploding offers
- Inefficiency: the matches are not stable and lead to aggregate losses
Unraveling: the market for gastroenterologists

- Before 1985, it suffered from the unraveling of appointment dates. A centralized match was implemented with: (1) a uniform appointment date and (2) applicants and fellowship programs submitting rank order lists.
- That form of organization, a centralized market with a clearinghouse, is widely believed to produce stable matching and are usually long lasting.
- After 1996, however, the centralized match broke down and was abandoned. Why?
The collapse of the GI clearinghouse

- The demise of the match seems to have been set in motion in 1993-4, when gastroenterology subjected itself to a manpower analysis which endorsed a sharp reduction of the supply for GI fellows.

- In 1996, for the first time, and despite the reduction in the number of positions offered, there were fewer applicants for GI fellowship positions than there were positions offered in the match.

- Despite the considerable reduction in the supply of positions, the market experienced a shortage in demand for positions, as residents stayed away from the market.
Why may a centralized market collapse?

Two asymmetries seem particularly important in the present instance:

▶ Asymmetry of actions: In both centralized and decentralized markets, applicants apply to firms to be considered for a position (e.g. apply for an interview). In decentralized markets, or when firms do not wait for a centralized market, firms then make offers, and applicants decide whether to accept or reject them.

▶ Asymmetry of information following a shock: Receipt of applications gives firms an informational advantage; they know if they are getting many or few applications, and hence they have an early indication of shifts in the demand for positions. In contrast, information about the supply of positions is common to both firms and applicants, since available positions are announced well in advance. And in the absence of a shock, historical information will be a reliable guide to both firms and applicants.
Why may a centralized market collapse?

Three related hypotheses about why the shortage of applicants for gastroenterology positions in 1996 set off the collapse of the match. Each is a special case of the previous hypothesis.

1. The success of this kind of match depends on there being more applicants than positions: the centralized match fails when there are fewer applicants.

2. The match failed because there was a shock that reduced the demand for positions below the supply, but the match could have recovered from this shock if given the chance, once supply and demand stabilized.

3. The match failed because there was a shock that reduced the demand for positions below the supply, and because firms knew this (because they could see their reduced applicant pool) and applicants didn’t. However, the match could have recovered from this shock if given the chance, once supply and demand stabilized, since then applicants would no longer be at an information disadvantage to firms.
Now what?

Recap:

1. It is widely believed that centralized market are efficient
2. The GI market is particular because it is one of the rare instance in which such a market collapsed (many profitable matches did not occurred)
3. Using economics insights and a careful description of the problem, we formulated some assumptions about the reasons of the collapse
4. So, now we would like to design an experiment that will help disentangle these different assumptions
5. HOW WOULD YOU DO THAT?
Experimental design

- A study of the effects of both stationary imbalances in supply and demand and of shocks that create a shortage of applicants.
- In each experimental condition, subjects first gain experience of unraveling by participating in 15 decentralized markets.
- They then participate in 15 markets in which centralized matching is available to subjects who choose to wait and use it.
- In the treatments that involve a shock, the shock occurs after the thirtieth market, and subjects participate in an additional 15 markets after the shock, with centralized matching available for those who wait to use it.
- Firms and applicants are always fully informed about the number and types of firms in the market.
- We vary the information that applicants have about other applicants:
  - full information: firms and applicants have the same information about supply and demand,
Participants objectives

- Participants in the experiment are assigned the role of either a firm or a worker of one of two types, high or low.
- Each subject maintained the same role.
- In each market, firms can match to at most one worker, and each worker can match to at most one firm.
- For each participant, a match to a High type is worth 150 points plus or minus a private value between 0 and 10 points. A match to a Low type is worth 50 points plus or minus a private value between 0 and 10 points.
The market

- In a decentralized market, firms have three periods to match by making offers to workers, who can accept or reject the offers.
- In a centralized market, the first two periods are as in the decentralized market. But in the last period, period 0, all firms and workers who have not already matched in periods -2 and -1 are matched at the firm-optimal stable matching.

The costs of unraveling:

- each participant who matches in period 2 incurs a cost of 20 points and
- each participant who matches in period 1 incurs a cost of 10 points.

The penalties for matching early are deducted from the participants earnings. Subjects who failed to match by the end of period 0 receive 0 points for that market.
The matching mechanism

- Markets are divided into three periods, and in each period, firms that are not yet matched can make up to one offer.
- After all the firms make offers (or choose not to make an offer), workers accept or reject offers.
- An offer contains the type and an identification of the firm that made that offer.
- A worker who receives several offers sees them only one at a time, in a random order.
- A worker must choose to accept or reject each offer without knowing whether he has any more offers coming in that period.
- Accepted offers are announced to all participants in any subsequent period, while unaccepted offers remain private.
- Matched firms are no longer allowed to make offers and unmatched firms cannot make offers to matched workers.
The treatments

- In each treatment there are 3 low type firms and 3 low type workers.
- The number of high type firms and workers to induce demand and supply imbalances varies.

Three fixed demand treatments: 30 rounds (15 decentralized then 15 centralized), full information:

1. 4 High type workers and 3 High type firms
2. 3 High type workers and 2 High type Firms
3. 2 High type workers and 3 High type Firms
The treatments

Two shock treatments, with full and partial information:

- Started out with 3 High type firms and 4 High type workers.
- After market 30, there was a single shock, eliminating 2 High type workers
- Resulting in a shortage of workers from markets 31 to 45.
- **Full Information**: Workers are fully informed about the number and types of workers (including the shocks)
- **Partial Information**: Workers are not informed about the number and types of workers in each market
Variable of interest

- That is we are concerned whether the maximal number of possible High-High matches are achieved in each market
Results: shocks and unbalanced supply and demand

- Each point represents the over five markets
- Within treatments differences use the Wilcoxon matched-pairs signed-rank test,
- Differences across treatment use Mann Whitney test
There is no suggestion in the data that centralized matches work well only when supply and demand are balanced, or only when firms are on the short side of the market.

In each of the experimental sessions with no shocks (or before the shock), the centralized match, once introduced, achieved a steady rate of participation, suggesting that it is robust to varying conditions of supply and demand.
Results: the effect of information

High type worker tend to accept their offers sooner rather than taking the risk of waiting. Remember that earlier matches are costly.
Results

- The same result (too many early matches) holds under partial information
- We get support for the third hypothesis (why?)
The three assumptions

Three related hypotheses about why the shortage of applicants for gastroenterology positions in 1996 set off.

1. The success of this kind of match depends on there being more applicants than positions: the centralized match fails when there are fewer applicants.

2. The match failed because there was a shock that reduced the demand for positions below the supply, but the match could have recovered from this shock if given the chance, once supply and demand stabilized.

3. The match failed because there was a shock that reduced the demand for positions below the supply, and because firms knew this (because they could see their reduced applicant pool) and applicants didn’t.
Results

Based on experimental evidence, the authors provides the following analysis:

- The match failed because there was a shock that reduced the demand for positions below the supply.
- Firms knew this and applicants did not.
- Recall how the low demand for positions in 1996 was an unexpected shock.
Defriefing: Unraveling

- What do you think about that experiment?
- In your view, is it an interesting experiment?
- What have you learned? What have we "proved"?
- What kind of follow-up study would you like to carry on?
- Is it important for economics (i.e. should we care?)
We looked at three type of experiments:

- The marshmallow experiment
- The Linda and Bob problems
- The GI market

What have you learned?

How representative are these experiment from the field?
A booming field

- The number of experimental papers published in top journals (source Card et al. JEP)
How big is the field?
What are the "classics" in experimental economics

95% of studies can be put into four categories
The big four

- **Social preferences (35% of all papers):**
  - Social dilemmas
  - Dictator ultimatum games
  - Trust, gift exchange

- **Individual decision making (14% of all papers):**
  - Risk (decision under uncertainty, etc)
  - WTA/WTP

- **Markets (24% of all papers):**
  - Auctions
  - Asset markets

- **Games (21% of all papers):**
  - Coordination
  - Beauty contests