WHY EXPERIMENT? An introduction to its purposes, history, and practices  
(Lecture 1, 2009_01_28)

Remark: The lecture notes are exactly that. They are my notes (mistakes and all) for 
the lecture. Nothing more. They are NOT a transcript of the lecture. (Rather they are 
my notes for the lecture but they will be, typically, revised after the lecture.)

Lecture notes will be posted after the lecture – normally on the same day - at
http://home.cerge-ei.cz/ortmann/TrentoCourse/TrentoCourse.html
There you also find the reading assignments (both required and optional), almost 
always with a direct link.

[Discussion of syllabus:
- Will revisit some of the assigned readings, sometimes repeatedly.
- You are expected to have read all required readings reasonably well.
- You are not expected to print out all readings.
(In fact, you are discouraged from doing so.)
- I consider pop quizzes and calling on people (incentive compatible) fair game
- No presentations; instead will do occasionally group work based on sets of guiding 
questions.
- We will often answer such sets of guiding questions in the plenum.
- There will be worksheets (typically online on the evening of the Wednesday lecture) 
that ought to give you some indications how well I expect you to know the assigned 
readings.
- The goal is to get you to the frontier of experimental economics research.]

EXPERIMENTAL ECONOMICS  (CIFREM, SECOND TERM) -  Version: Jan 28, 2009

Lecturer: Andreas Ortmann
E-mail: aortmann@yahoo.com, andreas.ortmann@cerge-ei.cz
Home page: home.cerge-ei.cz/ortmann, course materials available in due course at
.../ortmann/trentocourse/trentocourse.html; last year’s course materials (on which this year’s 
course is based) at .../ortmann/trentocouse2008/trentocourse2008.html.

Objective:

Over the past couple of decades experimental economics (EE) has become an indispensable, 
albeit not uncontroversial, tool in the economist’s toolbox. This course offers a crash course 
in the promises and pitfalls of the experimental approach in economics (and psychology); we 
will also attempt to clarify where and why behavioral economics and EE differ.

Reading materials:

There is no textbook for this course, as we will draw on a judicious mix of articles old and 
new. (I will, however, take the opportunity to test-drive a manuscript that is based on last 
year’s course.)

Tentative course outline:
Classes are typically on selected Mondays 9 - 11:30 and Wednesdays 11 _ 13 and start January 28, 2008 (a Wednesday).

L1 (Jan 28, W) Why experiment? The three purposes of EE. An introduction

L2 (Feb 2, M) Experiments and theory

L3 (Feb 4, W) Experiments and public policy: “Parallelism” and the need to calibrate

L4 (Feb 9, M) - Experiments and markets: Induced value theory and other tricks of testing theories.

L5 (Feb 11, W) - On the experimental practices of economists and psychologists 1: Lab environs, the “real world”, and the importance of context

L6 (Feb 16, M) - On the experimental practices of economists and psychologists 2: financial and/or social incentives?

L7 (Feb 18, W) - On the experimental practices of economists and psychologists 3: deception?

L8 (Feb 23, M) - On the experimental practices of economists and psychologists 4: semantics, pragmatics, and other issues

L9 (Feb 25, W) - On the importance of thinking about the statistical evaluation of experiments before you do them

L10 (March 4, W) - On subject pools and levels of reasoning

L11 (March 9, M) - On neuroeconomics

L12 (March 11, W) - Behavioral Economics and the controversy over the reality of cognitive illusions

L13 (March 23, M) - Exemplar: Overconfidence in economics - myth or reality?

L14 (March 25, W) - TBD (possibly contingent valuation studies, or time discounting issues, or coordination games)

Early April: Final (to be discussed in class, sometime between April 5 - 8)

You are expected to write a literature review (of between 5 - 7 experimental articles or papers) that ideally ought to, but does not have to, inform your search for a dissertation topic. The literature review is due on the day of the final. No exemption.

You may, alternatively, design an experiment. If you choose to take this option, you will have to write the draft of a paper that has to include the following sections: introduction/motivation, design/research hypotheses, implementation, and instructions. More details in class.
Grading:
Final exam 60 %
Lit Review/Draft of paper 30 %
Class participation 10 %

Office hours:
TBA
In general, I’m happy to talk any time (preferably over a good cup of coffee).
Contact me after class or via e-mail to make arrangements.

Historically, ...


Thurstone (1931) – testing experimentally the indifference curve representation of preferences (famously critiqued by Friedman & Wallis, 1942, for artificiality of experimental situation -> we'll return to this issue later) ... Allais Paradox and Ellsberg Paradox (Allais 1953; see also Kreps 1990) as examples of individual decision making under risk and uncertainty. ... theories of disappointment and regret as modifications of such theories ... See Starmer (2000) for an excellent discussion of the relevant experimental literature. (Discussed in Prof Mittone’s course?)

Chamberlain (1948; see also Smith 1962 and Smith 2008), Sauermeann & Selten (1959, 1960), Siegel & Fouraker (1960) and Fouraker & Siegel (1963) – testing experimentally the organization of markets and other IO issues ... Plott (see Ortmann 2003) extended the use of experimental methods to
- public economics, political processeses, and policy applications
- market institutions and the price discovery processes
- information, finance and general equilibrium

Dresher & Flood (in 1950) – testing experimentally game theoretic hypotheses (such as the Hangman’s Paradox, or PDG, a metaphor of arms races, public good provision, commons problems, price wars and so on – see Ortmann, JEE 2003, and literally thousands of other articles on it; see Roth’s Introduction to Kagel & Roth, 1995, pp. 9/10, for Nash’s critique of these experiments.) ... Nash bargaining solutions ... See Roth (1988) for an excellent discussion of the relevant experimental literature. See also Schelling (1957, 1960). Some of the current issues to be discussed in this course.

Growth data:
- fewer than 10 (30, 50) publications per year before 1965 (1975, 1985)
- more than 100 (200) publications since 1990 (1999)
- Economic Science Association since 1986

- Experimental Economics since 1998, since 2005 indexed, now has IF that establishes it soundly as second-tier journal

Today ... there is hardly an economics topic on which economists have not experimented ... (that includes macro topics, see http://www.pitt.edu/~jduffy/papers/hee11.pdf - John Duffy’s survey of laboratory research in macroeconomics)

See Davis & Holt (1993) - path-breaking book on experimental economics, still very readable. They claim as major insights EE has provided:

1. In many situations, neoclassical price theory explains observed behavior quite well.
2. Institutions matter.
3. Some predictions of game theory describe behavior well.
4. Other game-theoretic predictions have a more restricted range of application.
5. Even apart from the institutional specification, many results are characterized by a “gray” area where variables irrelevant to the theory affect outcomes.
6. Our understanding of individual behavior is incomplete; some recurrent anomalies are fundamental challenges to rational models of behavior. [Davis and Holt, 1993, 506-509]

See Kagel & Roth (1995)

  A masterful critical review asking all the right questions
- Ochs, Coordination Problems
- Roth, Bargaining Experiments
- Holt, Industrial Organization: A Survey of Laboratory Research
- Sunder, Experimental Asset Markets: A Survey
- Kagel, Auctions: A Survey of Experimental Research
- Camerer, Individual Decision Making
  Idiosyncratic, with a bia towards the heuristics-and-biases school

See Camerer (2003; see review by Ortmann & Rydval, JoEP 2004)

- Dictator, Ultimatum, and Trust Games
- Mixed Strategy Equilibria
- Bargaining
- Dominance-solvable Games (beauty-contests etc.)
- Learning
- Coordination
- Signaling and Reputation
- Top Ten Open Research Questions
  - How do people value the payoffs of others?
  - How do people learn?
  - How do social preferences vary across cultures?
  - What happens when people confront new games?
- How exactly are people thinking in games?
- What game do people think they are playing?
- Can experiments sharpen the design of new institutions?
- How do teams, groups, and firms play games?
- How do people behave in very complex games?
- How do socio-cognitive dimensions influence behavior in games?

See the current issues of journals ... (See also Smith 2008, Theory and Experiments: What are the questions? To be discussed in class soon)

See what working papers come out these days ...

Almost everyone has become a behavioralist economist (and dabbles in neuro-science).

Rational-choice theory gets attacked left and right (mostly by people whose knowledge of modern economic theory is ... wanting)

... the exciting developments in this field will be discussed in this course ...
also, how experimental economics is different from behavioral economics/finance/etc. and experimental psychology ...

Most interesting developments (from my point of view) ...

- Renewed reflection on design and implementation issues including the attempt to “leave the reservation“ (Harrison), i.e., doing field experiments
- The attempts of some to complement experiments with humans with those featuring computational agents
- The rational choice empire striking back (e.g., work by Cherry et al., Goeree & Holt, Harrison, List, Plott, Roth, etc.) – this very similar to the critique of the heuristics-and-biases school by people like Gigerenzer, Lopes, Krueger & Funder et al.

So, what are the uses of experimentation (today)?
(again, following Roth’s excellent Introduction to Kagel & Roth, 1995, pp. 3 – 109; see also Ortmann, JoEP 2003, on Plott’s work.)

- „Speaking to Theorists“ -> Lecture 2
  („theory-first“ approach to experimental research, as Plott calls it)
  - to test the predictions of well-articulated formal theories
  - to observe unpredicted „regularities“ that might lead to new theory formulations (see also Smith 1991)

- „Searching for Facts“ -> Lecture 2
  („data-first“ approach to experimental research, as Plott calls it.)
  - to study the effects of variables about which existing theory is moot (that might include subtle institutional differences and differences in design and implementation of experiments; says Camerer, „The way in which an experiment is conducted is unbelievably important.“ (2003, p. 34))
  - to understand causalities rather than correlations both in experimentation but also in theory
- “Whispering in the Ears of Princes” (“design economics”, as Roth has it, and before him Plott) -> Lecture 3
  - to answer questions raised by regulatory or other state agencies about the effect of changes in the way some market, or other economics activity is organized (see Klemperer 2002, Milgrom 2004, Roth 2002, and the recent flurry of literature on auctions and matching; see also earlier studies on airport landing slots, space station pricing policies, emissions trading mechanisms)
  - to study counterfactual designs at low costs (e.g., tests of particular auction formats or matching schemes)

- Getting right what (some) experimental psychologists get wrong. See assigned readings by Binmore, Grether & Plott, and Cherry et al., Bekkers, (and many others to follow)
Here is what we learn from Camerer (2003, section 2.1.) about the typical result of „dictator games“, or – maybe better – „reward allocation problems“. These are essentially ultimatum games where the responder does not have a chance to accept or reject. Hence, strictly speaking, it's not a game. Specifically, the „responder“ or „receiver“ is really a recipient.

“The fact that dictator offers are much lower than Proposer offers in ultimatum games, but positive, shows that Proposers are being both strategic (avoiding more to avoid rejection) and altruistic.“ (56)

Hoffman et al. (1994, 1996) showed that giving is, among other things, significantly affected by the „social distance“ between participants and experimenter.

Cherry et al. explore how „asset legitimacy“ (or, earned wealth, i.e., not „manna from heaven“ as is typical in experiments) interacts with social distance in reward allocation experiments.

Subjects were students at UCF.
Three treatments, each with more than 50 bargaining pairs that were randomly allocated to their „roles“ and rooms A and B (without being able to interact before, during, or after the experimental session).

Two stages: earning money and bargaining.

Money was earned by answering 17 GMAT questions under a time constraint.
- Those who answered at least 10 questions correctly, were paid $40.
- Those who answered less than 10 questions correctly, were paid $10.

Subjects were then separated into rooms A1 (high earners) and A2 (low earners).

Subjects in room A1 and A2 then decided upon the split of their earned wealth, i.e., their giving to an anonymous and randomly selected counterpart in room B.

Experimental earnings were paid according to the split.

The three treatments were baseline, earnings, and double blind with earnings. In all treatments there were those endowed with $10 and $40 dollars.

Baseline (B): „X has been provisionally allocated to each pair and the person in room A can propose how much of this each person is to receive.“

Earnings (E): „the person in room A has earned an amount of money by participating in a previous session“ and „the person in room B has not had the opportunity to earn any money.“ Also, „the person in room A decides how much of his or her earnings they are to receive and how much of his or her earnings the person in room B is to receive.“
Earnings under double-blind anonymity (DBE): Same as E but with bargaining segment under protocol also used by Hoffman et al. (1994, 1996)

Results? Well, what do you expect?

See Figures 1 and 2:
“other-regarding behavior is greatly diminished when bargaining involves earned wealth,”

- less than 20 percent (L: 19, H: 15) made zero offer in B session
- more than 70 percent (L: 79, H: 70) made zero offer in E session
- more than 90 percent (L: 95, H: 97) made zero offer in DBE session

All treatment effects are significant (Fisher exact test, Wilcoxon) except for the L treatment comparison of E and DBE.

Note the effect of financial incentives; it is not in all cases what economic theory predicts.

Conclusions:

1. Lack of asset legitimacy seems source of giving behavior in dictator games; how about other games (e.g., the alleged altruistic behavior in public good games, gift exchange games, etc.) ?

2. To revisit Camer’s claim: Is it altruism or strategic behavior that we see in Dictator, Ultimatum, and Trust games? This evidence seems to suggest it’s strategic behavior. (Demand effects? See also Guala & Mittone 2008))

3. How important is this result? (It’s only one of many after all! Or, is it?)

- It seems to overturn a long history of experimental results that suggest that people are not only strategic but altruistic. It seems to fundamentally challenge behavioralist challenges to rational choice theory („Cognitive Illusions Controversy“, to be discussed later in this course.
- It seems to have relevance for many other settings such as ultimatum, trust, gift exchange, principal-agent, public good provision etc.
- Mittone & Ploner (2008) – dictator behavior in Dictator game can be pushed in other direction in a Cherry et al (2002) setting by asking recipients to exert the same effort as the dictators (but not paying them) – up to 80% of dictators now willing to give something, the average level of donations is tripled compared to Cherry et al (2008) setting (asymmetric effort)

- Ruffle (1998) – dictator behavior in Dictator game can be pushed in other direction when recipients are asked to contribute to the size of the cake by answering a quiz. About 20% of dictators went as far as offering more than half of the sum that recipients had earned. Average level of donations also increased. (See also Engelmann & Strobel AER 2004, 2006)

- Eckel & Grossman (1996) – dictator behavior in Dictator game is affected by who the recipient is. A recipient such as the Red Cross will attract more donations. (And the level of donations is predictably affected by the way the organization spends the money: conjecture: the recent bad press that Red Cross had to endure / has triggered should reduce giving. Intentions matter. Altruism, in other words, is not a primitive.

Dictator game originally used in Kahneman, Knetsch, & Thaler (1986) „as a control treatment, i.e. as a set-up that helps investigating another game, which is the focus of their study.“ (p. 8) [the Ultimatum game – discuss]

„Public Goods and Ulitmatum Games elicit norms ... in virtue of their structure, and of the fact that games with such a structure are relatively common in many societies. The same, however, cannot be said of the DG. The DG has a remarkably simple structure – indeed too simple (we should perhaps say „poor“) to elicit a specific norm. It is an unusual situation too, for in real life one rarely deals with „windfall money“ to be shared with an anonymous stranger. It is important to stress both aspects: structure and familiarity. Subjects have to deal with windfall money and anonymous partners in Ultimatum games, but the structure of the UG is rich enough to focus their attention on power asymmetries, and thus elicit the fairness norms that apply in such circumstances. The DG in contrast is too „thin“ for that, and experimental subjects are left to puzzle over which behavior is deemed appropriate for a situation of that kind.“ (p. 11) [Really? – see Bekkers 2007, see also Samuelson 2005]

„We now know that behavior can both diverge and conform to the standard model, and we know that in different situations experimental subjects can be pushed towards one or the other extreme of the self-interested/other-regarding spectrum.“ (p. 14)

„ ... every experimental context is a social context of some kind: even the most ,purified design, where every element of sociality has been removed by the experimenter, must be interpreted by the experimental subjects (...). ... by removing every cue that may guide the players in forming normative expectationsm, we may end up with extremely fragile results that have little ecological validity.“ (p. 15)
### Why experiment in economics?
(Why not just leave the experimental testing to experimental psychologists?)

- because the „school of Kahneman and Tversky“ (also known as the heuristics-and-biases school) „has been criticised in the experimental psychology literature for the failure of their results to withstand simple robustness tests (Gigerenzer 1996). ... Under the circumstances in which the school of Kahneman and Tversky run experiments, we should not expect economic theory predicting well.“ [F16]

- [people are not omniscient human beings; „people get to equilibrium – insofar as they do – by an interactive process of trial-and-error learning. Taking this view requires accepting that they will not get to equilibrium at all if we put them in too complicated a situation, or fail to make paying attention to what is going on worth their while. ... economic theory should only be expected to predict in the laboratory if the following three criteria are satisfied:
  - The problem the subjects face is not only ‘reasonably’ simple in itself; but is framed so that it seems simple to the subjects;
  - The incentives provided are ‘adequate’;
  - The time allowed for trial-and-error adjustment is sufficient.
... Just as we need to use clean test tubes in chemistry experiments, so we need to get the laboratory conditions right when we test economic theory.“ [F17]

Do these criteria exclude field experiments? We'll address this issue later in detail. But note that Binmore speaks about the possibility of experiments being „a source of inspiration for revising the theory where it does not work so well“ [F17].

Can one always verify economic theory by interpreting the criteria „severely enough“? (as Binmore claims on F 18)
Can one always falsify economic theory by interpreting the criteria „loosely enough“? (as Binmore claims on F 18)
Two case studies:

- two-person, zero-sum games
design and implementation problems? „The only interesting question is whether people learn to play minimax in repeated trials against changing opponents.” [F19]
See also recent studies with real-world data (e.g.,
- Walker, Wooders (AER 2001), Minimax Play at Wimbledon.
- Palacios-Huerta (RES 2003), Professionals Play Minimax.
- Palacios-Huerta, Volij (Econometrica 2008), Experiencia Docet: Professionals Play Minimax in the Laboratory Experiments
- Levitt, List, Reiley (2007), What happens in the Field Stays in the Field: Professionals Do not Play Minimax in Laboratory Experiments
- Palacios-Huerta, Volij (2008), Field Centipedes

By and far, these data articles/papers suggest that professionals are pretty good at playing mixed strategies and doing backward induction. (Not at trivial result given much evidence to the contrary in psychology and economics.)

- the ultimatum game
design and implementation problems? „[l]t would be a good idea to use both experimental and theoretical tools to find out when and why backward induction predicts successfully and when and why it does not. ... We need a theory that successfully tracks the learning behavior of individuals.” [F21] Fat chance that! But see ... Johnson, Camerer, Sen, & Rymon (JET 2002), „Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining.“ (To be discussed later in this class.)

So why do people fail to backward induct in the ultimatum game (if they indeed do, Binmore would question that!)? (Binmore’s explanation may be found on pages F22. Or in his social contract books (1994, 1998, 2006), or his excellent Does Game Theory Work? The Bargaining Challenge (2007), or Rubinstein’s response-time study (2007); the latter to be discussed later in the course.)

But, we have have gotten ahead of ourselves (gamewise that is) ...
### Bekkers, Measuring Altruistic Behavior in Surveys: The All-or-Nothing Dictator Game

“Previous studies suffer from experimenter demands, often make unrealistic assumptions on asset legitimacy, and commonly use convenience samples of students as participants. These aspects of the context in which allocation decisions are made clearly influence the results ... “ (p. 139)

Data & Methods:

- 1,964 respondents of the Giving in the Netherlands Panel Survey (GINPS)
- collected May 2002
- Internet-based survey (no experimenter demand effects?)
- Participants answered survey and got paid for it, then ...
- Were given options for payment of their earnings:
  - in form of vouchers (national chains of department stores)
  - in form of „Air miles“
  - in form of a donation to one of three charitable causes:
    - Medecins Sans Frontiers
    - Aids Fund
    - Queen Wilhelmina Cancer Fund
- Importantly, it was either or choice
- On average participants had to allocate about Euro 9 (for 35 minutes of answering a questionnaire)
- Also studied various correlates (not of much interest for us here)

Results:

- 1,852 respondents (94.3 %) kept the reward (= chose vouchers or „Air miles“)
- of 112 subjects (5.7 %) that gave
  - Queen Wilhelmina Fund 63 donations (3.2 %)
  - Medecins Sans Frontiers 39 donations (2.0 %)
  - Aids Fund 10 donations (0.5 %)

- considerably less giving than in Eckel & Grossman (G&EB 1996) where 10.4% of participants gave their entire endowment to Red Cross
- the lower percentage not because random sample of Dutch population rather than „convenience sample“ of university students
- mostly standard demographic correlates (more education makes it more likely that a participant gives although being female makes it less likely, contrary to standard findings; curious result)
- giving to one of the three options correlates with self-reported philanthropy in the previous year
- differences in stakes do not influence donation decisions in the „game“ (we will return to this finding later in the course)

Discussion:
- Very interesting (and very clever, to my mind) example of how to effectively embed experimental “games” into household surveys

- Does asset legitimacy drive this result? (Likely.)

### Plott, Economics in 2090: The views of an Experimentalist. (See also Plott’s (SEJ 1991), Will Economics Become an Experimental Science?)

„The assumption is that within the next few decades the biases that now exist in the scientific research establishment will be removed and that the science of economics will be allowed to develop naturally along with the others.“ (88)

According to Plott progress to be made in

- microeconomic theory („remains a challenge“)
- political science
- organization science
- law
- repeated games, games of asymmetric information („will requires another several man-decades of experimental efforts (89)“

What about 2090?

Basic research:
- concerned with preference formation and change (including chemical, biochemical, and physiological influences)
- concerned with how people decide to decide (including problems of cognition)
- ... will be more specialized

Policy analysis:
- will be influenced by theory developments in mechanism design, public choice, social choice (e.g., Bolton & Dewatripont’s Contract Theory; see also Prendergast (JEL 1999))
- concerned with incentive compatibility
- ... will be „substantially complemented by the application of laboratory experimental techniques.“ (93)

Economics in the classroom
- „will be completely different. ... Students will learn about economic principles from their own experiences and not only from the experiences of other people as reported in books or pages of statistics. ... In spite of what will be known about individual choice behavior, the process of market convergence will still contain many mysteries in 2090.“ (93)

Holt (2007) seems to indicate that that might happen much earlier than 2090.