

Experimental Economics: What have we learned?
A Comment

Manuel Arellano

Barcelona GSE Economics "Trobada"
October 31, 2014

- I do not have the expertise to discuss what substantive lessons we learned from lab experiments.
- I have rather put together some comments about methodology from the perspective of a data-oriented outsider.
- They are intended to help bridge the gap between experimenters and empirical micro and macro economists.
- One lesson I learned is that the BGSE community has formidable strength in the field thanks to the important work of Jordi, Rosemary, Ramon and others.

1. Experiments are one of the ways we can identify treatment effects

- Building counterfactuals: Any evaluation involves a comparison of units who received treatment with units who did not. The question is which units best represent the treated units had they not been treated.
- As Harrison and List put it: "In some sense every empirical researcher is reporting the results of an experiment." Every researcher who treats a variable as exogenous effectively views their data as coming from an experiment.
- In an experimental setting this belief is built into the design of data collection.
- In other cases this belief is based on theory, auxiliary evidence or both.
- A conditional exogeneity assumption (as in matching methods) asserts that all variables that need to be adjusted for are observed by the researcher.
- In an instrumental variables setting unobserved controls are allowed as long as some covariates are independent of the potential outcomes.
- Testing these assumptions is difficult. Experimental methods are popular because of their potential for constructing clear-cut counterfactuals.

2. The issue of generalizability of results

- The best way of isolating the causal effect of interest is problem-specific.
- As in Falk and Heckman consider the response function:

$$Y = g(X, W)$$

Y is an outcome, X is the determinant of interest and W all other determinants.

- In a wage-effort study W includes demographics of participants, market institutions, order of moves, whether or not interactions are one-shot, etc.
- In general the level of Y response to X depends on the level of W and X .
- Many lab experiments have provided evidence for certain values of W .
- There may be trade-offs between the tightness of controls and the relevance of the values of the controlled variables.
- Tight control facilitates replicability, but the field offers a large range of variations in W , which may be relevant but hard to implement in the lab.
- Transporting experimental findings to new environments requires a model.

3. Methodological debates (and their decline)

- A good thing of modern economics is the increasing emphasis in substantive findings relative to methodology.
- The opportunities for complementarities between methodologies are now well understood, but realizing them requires less segmentation by research-style and retraining.

Experimental vs nonexperimental evaluation of treatment effects

- Debates between users of observational data and experimental data (sparked by Lalonde's 1986 critique).
- Rubin's matching methods: make non-experimental data look like experimental data (health effects of smoking).
- Notable examples that combine experimental data with theory and econometric methods: Ham & Lalonde (1996), Card & Hyslop (2005), and Todd & Wolpin (2006).

The role of theory: structural vs nonstructural approaches

- The current state of play in public finance and labor according to David Card:
- "When I started, most projects would have a pretty explicit theoretical front-end, and sometimes the best ones would then map that directly into the empirical approach.
- "Then in the '80s, it became less and less important to have this well worked-out theoretical framework. In some cases, people were focusing on extremely straightforward questions with much more emphasis on how carefully identified were the empirical results.
- "But in the last 10 years, there's been a backlash, and for almost all of my PhD students, I really emphasize the importance of having a well-posed theoretical model."

Lab versus field experiments vs RCTs

Lab

- The lab provides tight control variation, offering opportunities to control decision environments in ways that are hard to replicate in the field.
- Lab experiments are low cost relative to alternatives.
- Lab allows exogenous changes in institutions (market design, exchanges, regulations).

Lablike field

- Lablike field experiments differ from lab experiments in that they are conducted with nonstudent subject pools.
- Field experiments recruit subjects in the field and use field goods and field context motivated by a search for greater relevance for predicting field behavior, possibly at the expense of less tight control.

Lablike-field and randomized controlled trials (RCTs)

- Lablike field experiments differ from RCTs in that subjects make decisions that are not necessarily part of their day-to-day decision making, know that they are part of an experiment, or both.
- Lab-like data help understand mechanisms through which RCT-based treatment effects operate.
- For example, suppose an intervention is conducted in which different types of technologies are randomly offered to farmers and take-up is studied.
- Risk and time preferences elicited through lablike experiments can be used to test whether take up varies with these aspects.
- In an RCT the design often confounds several factors but the design is directly linked to an actual policy hopefully leading to reduced form causal treatment effects of policy relevance.

Need for considering the elicitation of truthful beliefs of researchers

- Schorfheide and Wolpin (2013) provide a formal rationale in the context of work that combines RCTs and structural estimation.
- Structural estimation provides the capability to extrapolate beyond the experimental treatment, but is subject to data mining.
- Data mining poses an impediment to the implementation of the first-best Bayesian analysis.
- They discuss a situation where combining the approaches by holding out from the estimation part of the sample allows for external validation.

4. Experiments as a data production technology for systematic use

- One of the main purposes of experiments is to measure (or elicit) individual characteristics that have traditionally been considered as "unobservable" (risk preferences, time preferences, propensities to trust or cooperate).
- These data can be used for purposes other than their primary one in combination with theory or additional data.

Harrison, Lau and Rutström's (2010) systematic approach

- They argue that rather than drawing inferences from individual experiments as if they were independent constructs, we need to constrain the inferences from one by the inferences from the other.
- Any data generated by an experiment needs to be interpreted jointly with considerations from theory and other data.
- They run large field experiments in Denmark to obtain measures of household and individual characteristics, such as risk preferences.
- They used econometric techniques to get estimates of risk attitudes that are consistent with estimated discount rates (Hey & Orme 1994, Camerer & Ho 1994).

5. Experiments on expectations and subjective expectations in macro

- An experimental literature that seeks to learn how subjects form predictions (surveyed in Duffy 2008; Assenza, Bao, Hommes & Massaro, 2014; and Wagener, 2014).
- Three types of experiments can be distinguished:
 - (i) experiments where agents predict an exogenously generated time series
 - (ii) Learning-to-forecast experiments (LtFEs, Marimon and Sunder).
 - (iii) Works that compare LtFEs with learning to optimize experiments.
- These studies are interested in understanding how agents form expectations.
- An alternative complementary approach is to get empirical survey data on agents' expectations (pioneered by Manski)
- Lab experimenters emphasized that surveys on expectations on future macro variables (like inflation) typically pay a fixed reward, which generate no incentive to provide a careful answer. But there are pros and cons.

Using subjective probabilistic questions to measure income risk

- Ask survey respondents about their subjective income expectations. Ask a question of the form:
- “What do you think is the percent chance that your total household income, before taxes, will be less than y over the next 12 months?”.
- The available evidence is that individuals are willing and able to respond to probabilistic questions about variables that are meaningful to them (Manski, 2004).
- Much progress has been made in understanding the implications of different methods of eliciting expectations.
- Data on subjective expectations and data on realizations can play complementary roles in constructing more credible measures of risk (Attanasio and Augsburg, 2012).

Using experiments within surveys to learn about survey measures of expectations

- Examples from the research program at the NY Fed around production and use of subjective expectations.
- Armantier et al (2011): they design an experiment within the survey to test whether households act on their subjective inflation expectations.
- Kuchler and Zafar: They find that the labor market situation of the respondent impact on their subjective expectations about the national labor market (eg chances that the unemployment rate would be higher).
 - Job searchers are less optimistic about future unemployment than employees.
- Basit Zafar et al: After asking for inflation expectations, a treated subset are provided information on experts expectations. Another subset are given information on past food inflation. Then they are asked about their expectations again.
 - Updating is higher in the case of those given information on experts forecasts.

Appendix A: Experimental Instruction (Ascending Scale Treatment)

You can earn extra money by answering the following 10 questions. In each question, you are asked to choose between 2 investments, investment **A** and investment **B**.

- If you choose investment **A**, then how much you earn depends on what the rate of inflation will be over the next 12 months. Your earnings under investment **A** depending on the rate of inflation are summarized in the table below:

Earnings under investment A												
Rate of inflation	-1% or less (deflation)	0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10% or more
Earnings	\$600	\$550	\$500	\$450	\$400	\$350	\$300	\$250	\$200	\$150	\$100	\$50

For example, we can see in the table that your earnings under investment **A** will be \$50 if the rate of inflation over the next 12-months is 10% or more. Alternatively, your earnings under investment **A** will be \$600 if the rate of inflation over the next 12-months is -1% or less (deflation).

- If you choose investment **B**, then how much you earn will not depend on the rate of inflation. Exactly how much you earn under investment **B** will be specified in each of the 10 questions below.

Once the survey is completed, we will randomly pick 1 of the 10 questions, and 2 survey participants. Twelve months from now, these 2 participants will be paid extra money according to the investment choice they made for the selected question. So answer every question carefully, as you may earn up to several hundred dollars. For investment **A**, the inflation rate over the next 12 months will be based on the official U.S. CPI index (Consumer Price Index) and it will be rounded to the nearest percentage point.

Earnings under investment **A**

Rate of inflation	-1% or less (deflation)	0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10% or more
Earnings	\$600	\$550	\$500	\$450	\$400	\$350	\$300	\$250	\$200	\$150	\$100	\$50

For every question, please choose between investment **A** and investment **B**.

Question 1: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$100**?

Question 2: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$150**?

Question 3: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$200**?

Question 4: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$250**?

Question 5: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$300**?

Question 6: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$350**?

Question 7: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$400**?

Question 8: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$450**?

Question 9: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$500**?

Question 10: Which one of these two investments do you choose?

- Investment **A**: your earnings are determined by the table above.
- Investment **B**, your earnings are exactly **\$550**?

Other experiments within surveys

- An early one: Barsky, Juster, Kimball, and Shapiro (1997) report measures of risk tolerance, time preference, and intertemporal substitution. These measures are based on survey responses to hypothetical situations in the HRS.
- Card, Mas, Moretti, and Saez (2012, AER): Inequality at work: the effect of peer salaries on job satisfaction.
- Online survey platforms: Kuziemko, Norton, Saez, and Stantcheva (2013) use randomized online surveys on Amazon Mechanical Turk (mTurk).
 - They look at the effect of information on policy views (on increased support for the estate tax in this instance)

Experiments in the public administration

- An illustration: Hesselius, Johansson and Larsson (2005) exploit a social experiment carried out in 1988 in Sweden to identify the effect of monitoring on sickness absence.
- The treatment consists of postponing the first formal point of monitoring during a sickness absence spell, a requirement for a doctor's certificate, from day 8 to day 15.
- The experiment was conducted in Gothenburg and Jämtland, and the treatment group was randomized by birth date.

6. Ethics in Economics

- Learned societies such as the ES and the AEA discussed the issue of ethical standards for papers that involve experiments on human subjects.
- US institutions are required to have an Institutional Review Board (IRB) on human participants.
- An IRB is a committee that has been formally designated to approve, monitor, and review biomedical and behavioral research involving humans.
- The ES discussed how to deal with experimental papers that come from institutions outside the US that do not require an IRB procedure. It was agreed to take the IRB standard as the norm, but implementation may not be easy.
- The EEA set up a committee on ethics last year to report on the practices among economists in Europe in several dimensions: (i) data collection; (ii) data management; (iii) ethical issues for RCTs and (iv) ethic for lab experiment.
- Rachel Glennerster (2014) provides some discussion of ethic issues in RCTs and practical tips for ensuring compliance with IRB.