Design, Environment and Institution

Wang, Yu

dr.yuwang@outlook.com

December 24, 2024

Overview

1. Experiment Design
2. Group Size and Assignment
3. Endowment, Payment and Stake
4. Decision and Behavior

Definition of Economic Experiments

Smith (1982)

- 1) Environment

  Number of subjects, endowments, preferences, costs, knowledge, and skills, and the economic constraints ... induced by appropriate (typically monetary) incentives linked to performance (no flat payment);

- 2) Institution

  Rules of the game, action set, timing of actions, payoff determination that define what subjects can and cannot do;

Definition of Economic Experiments

- 3) Behavior

  Given their environmental circumstances, subjects express their behavior by sending "messages" that the institution permits (e.g., bids, asks, acceptances, purchases, etc).

- 4) Outcome

  Via the governing rules, the institution maps subjects' choice of "messages" into experimental outcomes.

After framing the question of interest, the experimenter designs the environment and institutions to answer that question.

Definition of Economic Experiments

List (2007)

- An economic environment consists of a set of agents $(1, ..., n)$ and commodities $(1, ..., k)$.
- Each agent is described by a utility function, u_i, a technology or knowledge endowment, K_i, and a commodity endowment, ω_i.
- Each agent is therefore described by $\epsilon_i(u_i, K_i, \omega_i)$, and the microeconomic environment is defined by the collection of agents, $\epsilon = (\epsilon_i, ..., \epsilon_n)$.

Definition of Economic Experiments

- To complete the microeconomic environment, the experimenter specifies the institutional setting, I.
- The institutional setting includes the message space, M, the allocation rules, H, and other relevant characteristics of the specific institution of interest.
- The experimental system, $S = (\epsilon, I)$, thus is composed of the microeconomic environment and the institution.
- Agents, who are assumed to have consistent preferences and to be utility maximizer, choose messages and the institution that determines allocations via the governing rules.

Table of Contents

Steps of Experiments

- 1) Motivate research question and objective of the experiment.

  To test a theory, investigate anomaly, etc.
- 2) Design experiment and materials, including instructions, program, etc.

  Instructions should be as clear as possible but not lead subjects.
- 3) Seek and obtain approval from the Institutional Review Board (IRB).
- 4) Conduct experiment and pay subjects at the end individually based on their earnings in the experiment.

  Possibly post-experimental questionnaire to better understand what subjects did and why.
- 5) Analyze data.

  It is important to think about this step when designing the experiment, no data mining, report results honestly.

Treatment and Control

Experiments often involve comparing outcomes across different treatments.

- Treatment: a set of conditions subjects are being observed under;
- Treatment group: a group of subjects who receives a treatment;
- Control (treatment) group: a group of subjects who receives a standard treatment, a placebo, or no treatment at all.

Confounding

Confounding factor

- A variable that influences both the dependent variable and independent variable, causing a spurious association;
- Uncontrolled confounding factor can cause endogeneity and lead to biased conclusion.

Three tools to control the assignment of subjects to treatments and the influence of confounding factors:

- 1) Randomization;
- 2) Replication;
- 3) Blocking.

Randomization

If we randomly assign subjects to treatments, confounding factors still exist but affect one treatment as much as another.

- Random recruiting: we randomly choose whom to be invited to experiments;
- Random orders: we run experiment trials in random order (if possible);
- Random locations: if we run the experiment at multiple places, we randomly decide what treatments to run.

Replication

The average of a number of observations will be less noisy than a single observation.

- In a prisoner dilemma game (PDG), if we only have one single observation, we cannot draw a conclusion whether subjects tend to play “Defect” or “Cooperate;”
- If we have 100 subjects playing PDG, we will find that subjects tend to play “Cooperate.”

Replication tends to average out idiosyncrasies due to noise that can ruin statistical inference.

Blocking

We identify a "block" that is likely to have **the same confounding factor**, and then run each of treatments on that block. This reduces sources of variability and thus leads to greater precision.

- Gender: an experiment is designed to test a new drug on patients. There are two levels of the treatment, drug, and placebo, administered to male and female patients. The sex of the patient is a blocking factor accounting for treatment variability between males and females.
- Elevation: an experiment is designed to test the effects of a new pesticide on grass. A pesticide treatment group and a placebo group are applied to both the high elevation and low elevation areas of grass. It is blocking the elevation factor which may account for variability in the pesticide's application.

Blocking

- Completely randomized design

  Randomly assign subjects and everything else to treatments. The confounding factor is **unknown** and uncontrolled.

- Randomized complete block design

  Find subsets of the subjects that are **homogeneous**, and all treatments are run on each of these homogeneous blocks. The confounding factor has to be systematic and **known**.

Within Design vs. Between Design

Within design is a common form of blocking.

- Within design: treatments are run on **the same subject** or same group of subjects.

  There will be less data and fewer sessions required. But there might be learning or other contamination that affects observations on treatments.

- Between design: no subject or group of subjects experiences the same treatment twice.

  There is total separation between treatments. But there will be more data and sessions required.

One-Factor-at-a-Time Design vs. Factorial Design

One-factor-at-a-time design: to test how two (or more) factors affect outcome, we vary each one factor **independently**.

- For example, if we would like to study how (1) group size and (2) multiplication factor affect the players' contributions in public goods games (PGG).
- The 1st control treatment: group size is 8 and multiplication factor is 1.5;
- The 2nd treatment to study the effect of group size: group size is 12 and multiplication factor is ?;
- The 3rd treatment to study the effect of multiplication factor: group size is ? and multiplication factor is 2.0.

One-factor-at-a-time design ignores the **interaction effects** between factors (i.e., the interaction term "?").

One-Factor-at-a-Time Design vs. Factorial Design

One-factor-at-a-time design: to test how two (or more) factors affect outcome, we vary each one factor **independently**.

- For example, if we would like to study how (1) group size and (2) multiplication factor affect the players' contributions in public goods games (PGG).
- The 1st control treatment: group size is 8 and multiplication factor is 1.5;
- The 2nd treatment to study the effect of group size: group size is 12 and multiplication factor is 1.5;
- The 3rd treatment to study the effect of multiplication factor: group size is 8 and multiplication factor is 2.0.

One-factor-at-a-time design ignores the **interaction effects** between factors (i.e., the interaction term "group size X multiplication factor").

One-Factor-at-a-Time Design vs. Factorial Design

Factorial design: to test how two (or more) factors affect outcome, we observe each level of each factor along with each level of each other factor.

- For example, if we would like to study how (1) group size and (2) multiplication factor affect the players' contributions in public goods games (PGG).
- The 1st control treatment...;
- The 2nd treatment...;
- The 3rd treatment...;
- To study the interaction effect of group size and multiplication factor, the 4th treatment: group size is ? and multiplication factor is ?.

Factorial design incorporates the **interaction effects** between factors.

One-Factor-at-a-Time Design vs. Factorial Design

Factorial design: to test how two (or more) factors affect outcome, we observe each level of each factor along with each level of each other factor.

- For example, if we would like to study how (1) group size and (2) multiplication factor affect the players' contributions in public goods games (PGG).
- The 1st control treatment...;
- The 2nd treatment...;
- The 3rd treatment...;
- To study the interaction effect of group size and multiplication factor, the 4th treatment: group size is 12 and multiplication factor is 2.0.

Factorial design incorporates the **interaction effects** between factors.

Full Factorial Design vs. Fractional Factorial Design

Full factorial design

- There are **higher order of interactions** when the number of factors is larger than two.

  e.g., if we would like to study how (1) group size, (2) multiplication factor, and (3) the existence of punishment affect the players' contributions in PGG, there will be an interaction term of "group size X multiplication factor X dummy of punishment."
- With N factors each with k levels, there will be totally ? treatments.

  e.g., in full factorial design with 5 factors each with 2 levels, there will be totally ? treatments.
- Generally these higher order interactions have little effect and are difficult to interpret.

Full Factorial Design vs. Fractional Factorial Design

Full factorial design

- There are **higher order of interactions** when the number of factors is larger than two.

  e.g., if we would like to study how (1) group size, (2) multiplication factor, and (3) the existence of punishment affect the players' contributions in PGG, there will be an interaction term of "group size X multiplication factor X dummy of punishment."
- With N factors each with k levels, there will be totally k^N treatments.

  e.g., in full factorial design with 5 factors each with 2 levels, there will be totally $2^5 = 32$ treatments.
- Generally these higher order interactions have little effect and are difficult to interpret.

Full Factorial Design vs. Fractional Factorial Design

Fractional factorial design

- We only capture interactions between each pair of factors, but not interactions between interactions;

  i.e., no interaction term of “group size X multiplication factor X dummy of punishment.”
- The underlying assumption is that the higher order interactions have no effect;
- It can reduce the number of sessions to run.

Table of Contents

Group Size

Isaac, R. Mark, James M. Walker, and Arlington W. Williams. "Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups." Journal of Public Economics 54, no. 1 (1994): 1-36.

- Large group provided the public good more efficiently than small group.

Franzen, Axel. "Group size and one-shot collective action." Rationality and Society 7, no. 2 (1995): 183-200.

- There are no group-size effects in the one-shot Prisoner's dilemma game and Chicken game. However, group size does have a negative effect on cooperation rate in both the Volunteer's Dilemma and in the Assurance game.

Group Size

Barcelo, Hélene, and Valerio Capraro. “Group size effect on cooperation in one-shot social dilemmas.” Scientific Reports 5, no. 1 (2015): 1-8.

- Larger groups are more cooperative in the Public Goods game, but less cooperative in the N-person Prisoner’s dilemma.

Isaac, R. Mark, and James M. Walker. “Group size effects in public goods provision: The voluntary contributions mechanism.” The Quarterly Journal of Economics 103, no. 1 (1988): 179-199.

- Increasing group size leads to a reduction in allocative efficiency when accompanied by a decrease in marginal return from the public good.

Matching Protocol

Fixed matching protocol: subjects always interact with the same person and generally support a significant level of (sometimes full) cooperation.

- Palfrey, Thomas R., and Howard Rosenthal. “Repeated play, cooperation and coordination: An experimental study.” The Review of Economic Studies 61, no. 3 (1994): 545-565.
- Aoyagi, Masaki, and Guillaume Fréchette. “Collusion as public monitoring becomes noisy: Experimental evidence.” Journal of Economic theory 144, no. 3 (2009): 1135-1165.
- Bó, Pedro Dal. “Cooperation under the shadow of the future: experimental evidence from infinitely repeated games.” American economic review 95, no. 5 (2005): 1591-1604.

Matching Protocol

Random matching protocol: in any given period subjects meet in pairs but after each period matches are destroyed and new pairs are formed drawing subjects at random from the entire economy.

- Camera, Gabriele, and Marco Casari. "Cooperation among strangers under the shadow of the future." American Economic Review 99, no. 3 (2009): 979-1005.
- Schwartz, Steven T., Richard A. Young, and Kristina Zvinakis. "Reputation without repeated interaction: A role for public disclosures." Review of Accounting studies 5, no. 4 (2000): 351-375.
- Duffy, John, and Jack Ochs. "Cooperative behavior and the frequency of social interaction." Games and Economic Behavior 66, no. 2 (2009): 785-812.

Duffy and Ochs (2009) found remarkably higher cooperation in fixed than in random matching economies.

Group Identity

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir. “Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study.” The American economic review (1991): 1068-1095.

- Exogenous group identity generated by their platoons in the Swiss Army.

Goette, Lorenz, David Huffman, and Stephan Meier. “The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups.” American Economic Review 96, no. 2 (2006): 212-216.

- Exogenous group identity generated by their platoons in the Swiss Army.

Group Identity

Chen, Yan, and Sherry Xin Li. "Group identity and social preferences." American Economic Review 99, no. 1 (2009): 431-57.

- Endogenous group identity by painting preference.

In social preference games, subjects are more prosocial (more charity, less envy etc.) towards an ingroup match than that towards an outgroup match.

Table of Contents

Endowment

Cherry, Todd L., Stephan Kroll, and Jason F. Shogren. "The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab." Journal of Economic Behavior & Organization 57, no. 3 (2005): 357-365.

- Four-subject public good game with windfall or earned endowments;
- In the treatment, subjects take a quiz containing GMAT questions;
- Contribution levels were significantly lower when groups had heterogeneous rather than homogeneous endowments.

Subjects may exhibit different levels of social and redistributive preferences if their endowments are earned, rather than windfall.

Endowment

Konow, James. "Fair shares: Accountability and cognitive dissonance in allocation decisions." American Economic Review 90, no. 4 (2000): 1072-1091.

- Preparing letters for mailing.

Rutström, E. Elisabet, and Melonie B. Williams. "Entitlements and fairness:: an experimental study of distributive preferences." Journal of Economic Behavior & Organization 43, no. 1 (2000): 75-89.

- Solving a computerized puzzle.

Fahr, René, and Bernd Irlenbusch. "Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment." Economics Letters 66, no. 3 (2000): 275-282.

- Cracking walnuts.

Payment Scheme

Three payment schemes in infinite-horizon prisoner's dilemma:

- Cumulatively for all periods of the game;
- For the last period only;
- For one of the periods, chosen randomly.

Cooperation rates are not significantly different under the cumulative and the last period payment schemes, but they are significantly lower under the random payment scheme.

Stake Size

Subjects generally do not markedly change their behavior as the stakes increase.

- Slonim, Robert, and Alvin E. Roth. "Learning in high stakes ultimatum games: An experiment in the Slovak Republic." Econometrica (1998): 569-596.
- Cameron, Lisa A. "Raising the stakes in the ultimatum game: Experimental evidence from Indonesia." Economic Inquiry 37, no. 1 (1999): 47-59.
- Munier, Bertrand, and Costin Zaharia. "High stakes and acceptance behavior in ultimatum bargaining." Theory and Decision 53, no. 3 (2002): 187-207.
- Carpenter, Jeffrey, Eric Verhoogen, and Stephen Burks. "The effect of stakes in distribution experiments." Economics Letters 86, no. 3 (2005): 393-398.

Stake Size

Andersen, Steffen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A. List. "Stakes matter in ultimatum games." American Economic Review 101, no. 7 (2011): 3427-39.

- The range of endowment is from 20 rupees to 20,000 rupees (1.6 hours to 1,600 hours of work).
- For proposers, while the offer proportions are significantly lower in the higher stakes treatments compared to the lowest stakes treatment, the actual amount offered increases as stakes increase.
- For responders, at low stakes we observe rejections in the range of the literature, in the highest stakes condition we observe only a single rejection out of 24 responders.

Table of Contents

Action Space

The 11–20 game:

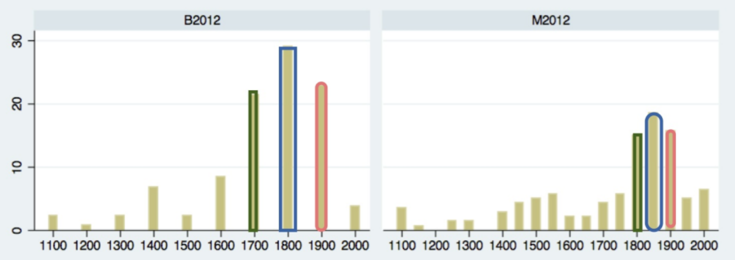
- "You and another player are playing a game in which each player requests an amount of money. The amount must be (an integer) between \$11 and \$20. Each player will receive the amount that he requests. A player will receive an additional amount of \$20 if he asks for exactly one shekel less than the other player. What amount of money would you request?"

Action Space

Three variants of the 11–20 game with different strategy spaces (Choo, Lawrence CY, and Todd R. Kaplan. "Explaining Behavior in the "11-20" Game." (2014)):

- Baseline (B2012) Game: {2000, 1900, 1800, 1700, ..., 1100}
- Medium (M2012) Game: {2000, 1950, 1900, 1850, ..., 1100}

In addition to the chosen payoff, a subject receives the bonus 2000 if the subject's amount is 100 Less than the other player's amount.



Decision Time

There is a close connection between short response time and choices that are clearly a mistake.

- Rubinstein, Ariel. "Instinctive and cognitive reasoning: A study of response times." The Economic Journal 117, no. 523 (2007): 1243-1259.
- Rubinstein, Ariel. "Response time and decision making: An experimental study." Judgment and Decision Making 8, no. 5 (2013).

He collected data on response time in web-based experiments on his website http://gametheory.tau.ac.il.

- Teachers who register on the site can assign problems to their students from a bank of decision and game situations.
- The students are promised that their responses will remain anonymous.
- The tasks are in non-laboratory setting without monetary incentives.

Lab, Field or Online

The external validity of experiments is whether the same individuals act in experiments as they would in the field.

- Benz, Matthias, and Stephan Meier. "Do people behave in experiments as in the field?—evidence from donations." Experimental economics 11, no. 3 (2008): 268-281.
- Frey, Bruno S., and Stephan Meier. "Pro-social behavior in a natural setting." Journal of Economic Behavior & Organization 54, no. 1 (2004): 65-88.

Pro-social behavior is more accentuated in the lab, and pro-social behavior in experiments is correlated with behavior in the field.

Lab, Field or Online

Online experiments have weaker experimental controls and less motivation compared to laboratory-based ones.

- Greiner, Ben. "An online recruitment system for economic experiments." (2004): 79-93.
- Hergueux, Jérôme, and Nicolas Jacquemet. "Social preferences in the online laboratory: a randomized experiment." Experimental Economics 18, no. 2 (2015): 251-283.

References I

[1] Kohei Nitta. *Experimental Issues: Influences of the Environment and the Design Parameters*. 2015.

[2] Kohei Nitta. *Experimental Issues: Payment Methods*. 2015.

[3] Kohei Nitta. *Experimental Techniques and Methods: Subject/Participation Selection*. 2015.

[4] Ryan Oprea. *Experimental Design*. 2014.