Introduction to Economic Experiments

Wang, Yu

dr.yuwang@outlook.com

December 24, 2024

Overview

Table of Contents

Motivations of Experiments

What are **experiments**?

Motivations of Experiments

Experiment is a scientific procedure undertaken to make a discovery, test a hypothesis, or demonstrate a known fact.

- 1) Test a theoretical claim;

  e.g., to test the Law of One Price.
- 2) Study an underlying behavioral mechanism;

  e.g., to study the cause of endowment effect.
- 3) Whistleblow to a real-world institution;

  e.g., to give advice to central bank policymakers.
- 4) Measure a behavioral parameter;

  e.g., to measure discount factor δ in time preference model.

Terminology of Experiments

- 1) Subject: a single human participant we are doing observations on;
- 2) Treatment: a set of conditions subjects are being observed under;

  e.g., in a study of the effect of vitamin A supplements on infant mortality in rural China, subjects in the treatment are observed under the condition that they take specific amount of vitamin A supplements.
- 3) Experiment: a collection of ? on subjects meant to answer a question;
- 4) Session: a single meeting at which observations are made on a group of subjects;

Terminology of Experiments

- 1) Subject: a single human participant we are doing observations on;
- 2) Treatment: a set of conditions subjects are being observed under;

  e.g., in a study of the effect of vitamin A supplements on infant mortality in rural China, subjects in the treatment are observed under the condition that they take specific amount of vitamin A supplements.
- 3) Experiment: a collection of **treatments on subjects** meant to answer a question;
- 4) Session: a single meeting at which observations are made on a group of subjects;

Terminology of Experiments

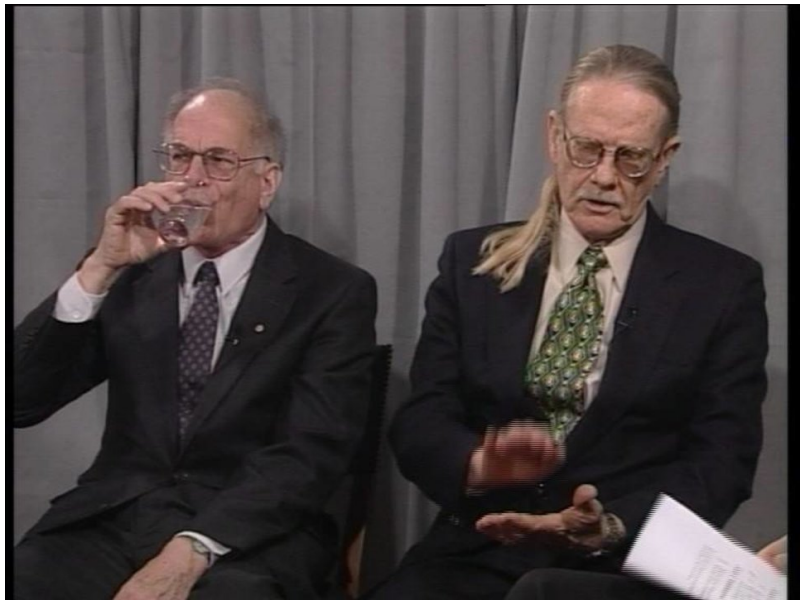
- 5) Factor: a variable hypothesized to affect the outcome of an experiment;

  e.g., the group size in a public goods game (PGG), the multiplication factor in PGG, whether PGG is a repeated game or one-shot.
- 6) Level: a setting for a factor.

  e.g., large or small, high or low, yes or no.

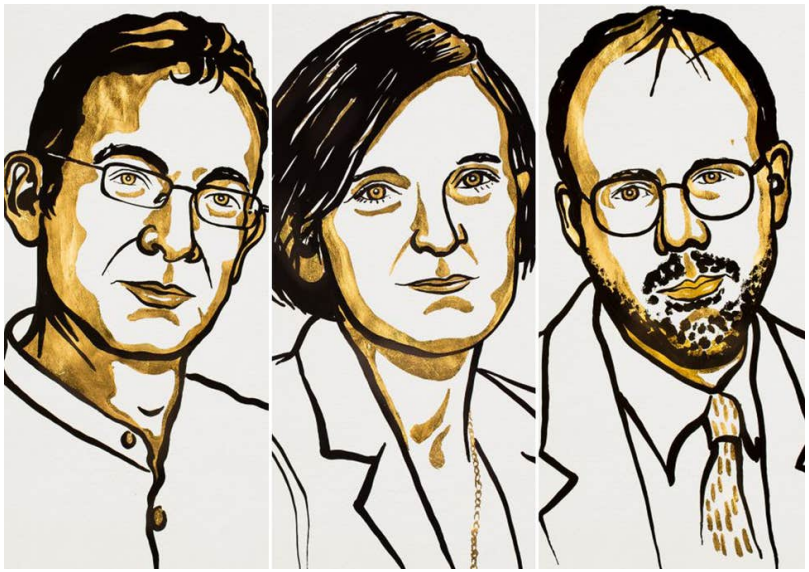
Controlled Laboratory Experiments

- An investigation carried out in a small-scale environment **in the laboratory** where adequate control is maintained.
- List (2007): *"Lab experimentation is the most convincing method of creating the* counterfactual, *since it directly constructs a control group via randomization."*



Field Experiments

- An investigation carried out **in the natural environment** outside of laboratory settings.
- List (2007): *"... however, field experiment occurs in the natural environment of the agent being observed and cannot be reasonably distinguished from the tasks the agent has entered the marketplace to complete."*



Field Experiments

Three subcategories of field experiments (List, 2007)

- 1) Artefactual field experiments

  The same as conventional lab experiments but with a **non-standard subject pool** (i.e., non-student sample);

- 2) Framed field experiments

  The same as artefactual field experiments but with **field context** in either the commodity, task, or information set that the subjects use;

- 3) Natural field experiments

  The same as framed field experiments except that the **subjects do not know** that they are participants in an experiment.

Field Experiments

A randomized controlled trial (RCT) is a type of scientific experiment (usually field experiment in economics) that aims to reduce certain sources of bias when testing the effectiveness of new treatments.

- The assignment of different units to different treatment groups is chosen randomly;
- This ensures that no unobservable characteristics of the units are reflected in the assignment;
- Any difference between treatment and control units reflects the impact of the treatment;
- RCTs can answer the questions that are causal in nature:

  e.g., what would be the impact of class size on students' learning outcomes?

Nobel Prize to Beh/Exp Economists

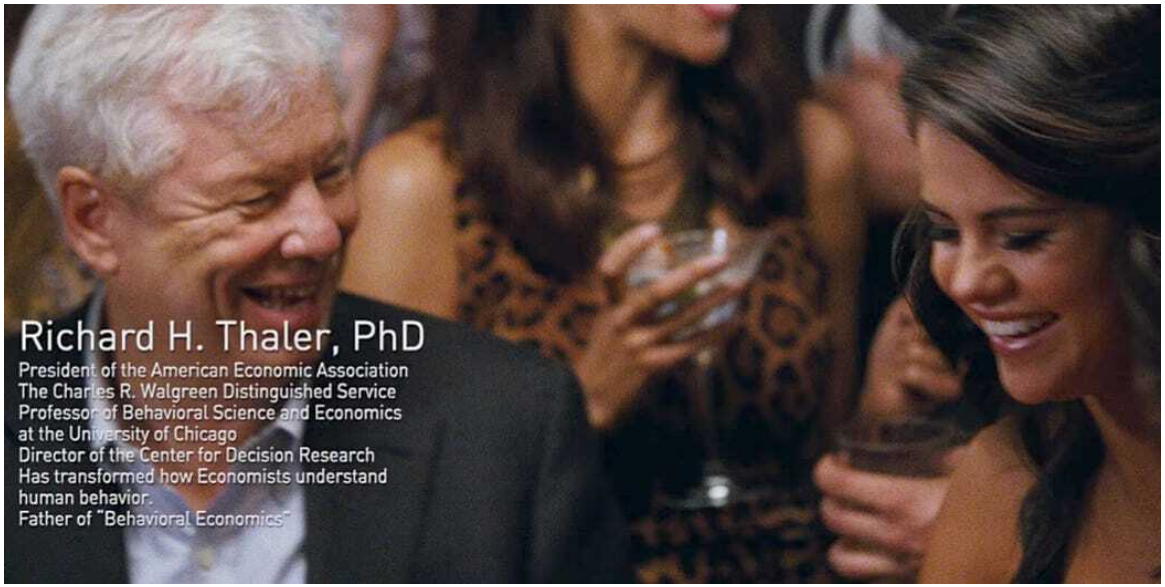
- 1978, Herbert Simon (bounded rationality, satisficing)
- 1992, Gary Becker (racial discrimination, altruism, rational addiction)
- 1994, Reinhard Selten (bounded rationality)
- 2001, George Akerlof (identity economics)
- 2002, Daniel Kahneman (prospect theory, heuristics and biases, etc.)
- 2002, Vernon Smith (laboratory experiment, market mechanism)
- 2012, Alvin Roth (game theory, market design)
- 2013, Robert Shiller (behavioral finance)
- 2014, Jean Tirole (motivated cognition)
- ...

Nobel Prize to Beh/Exp Economists

- 2017, Richard Thaler

  "for his contributions to behavioural economics (e.g., ?)."
- 2019, Abhijit Banerjee, Esther Duflo & Michael Kremer

  "for their experimental approach (i.e., ?) to alleviating global poverty."

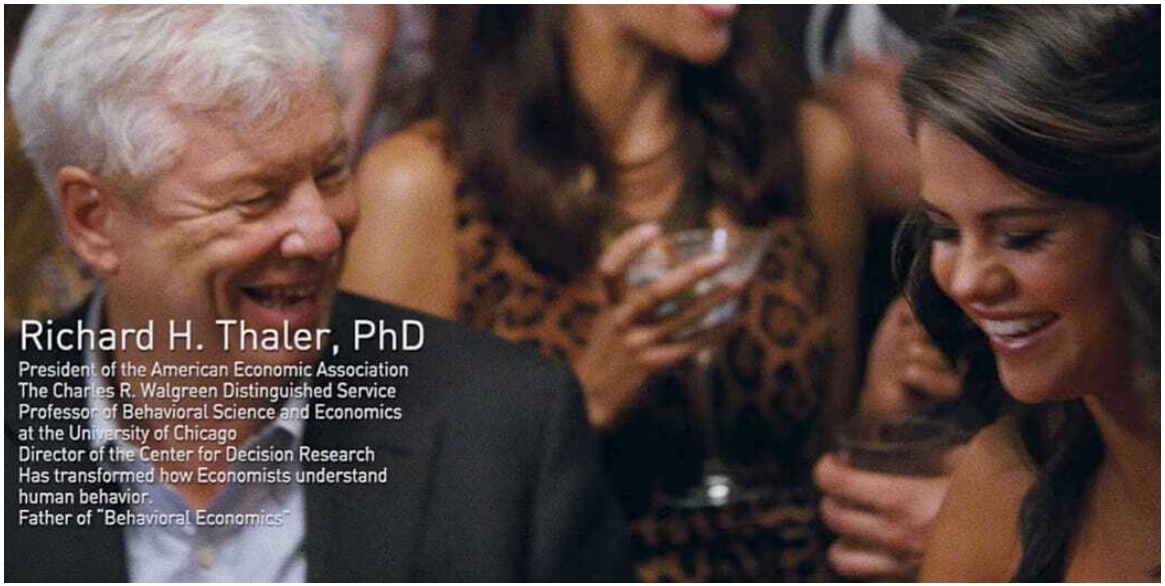


Nobel Prize to Beh/Exp Economists

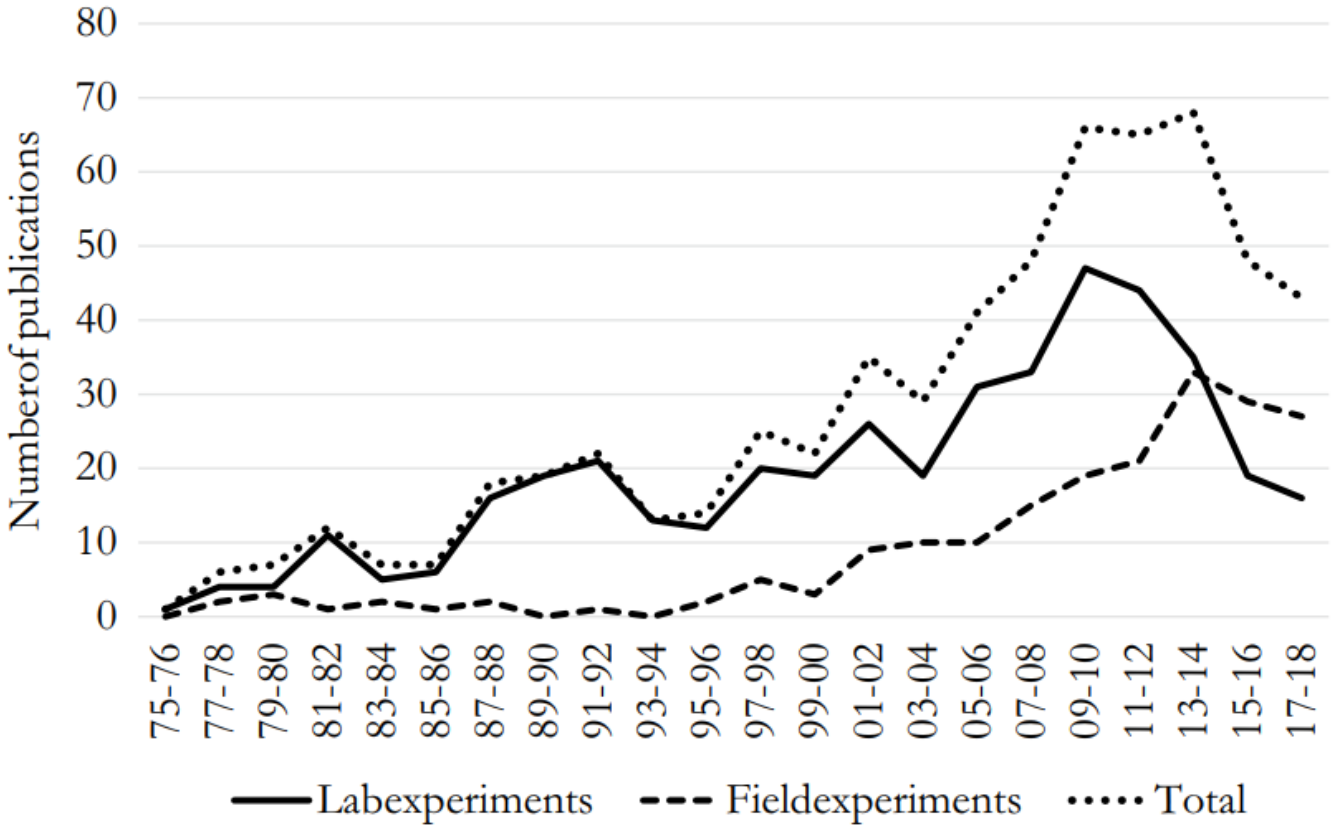
- 2017, Richard Thaler

  "for his contributions to behavioural economics (e.g., nudge theory, mental accounting, endowment effect, choice architecture, etc.)."
- 2019, Abhijit Banerjee, Esther Duflo & Michael Kremer

  "for their experimental approach (i.e., RCTs/field experiments) to alleviating global poverty."



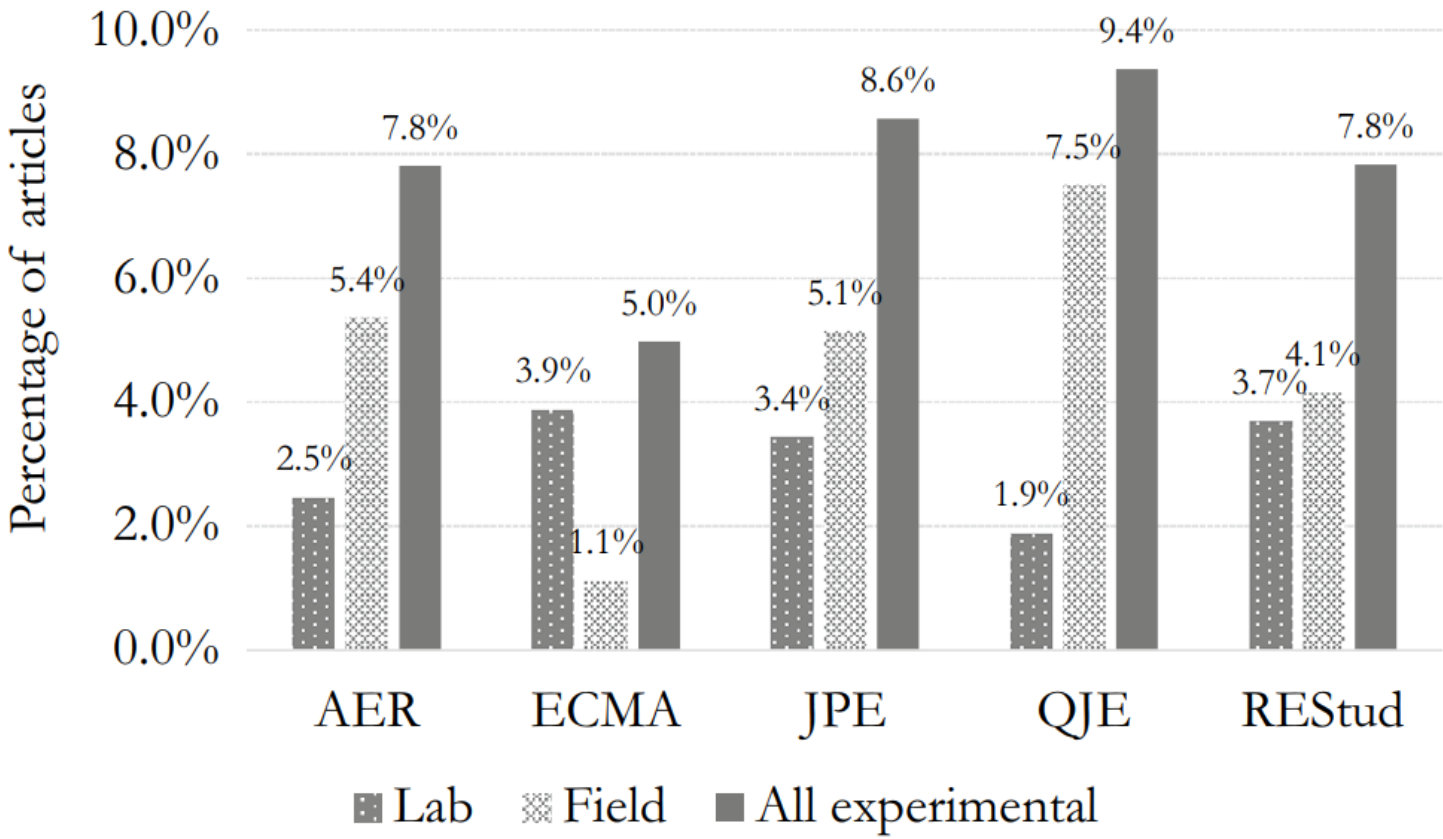
Recent Trends



Experimental papers published in the "top-5" economics journals (1975–2018)

Figure: Nikiforakia and Slonim (2019)

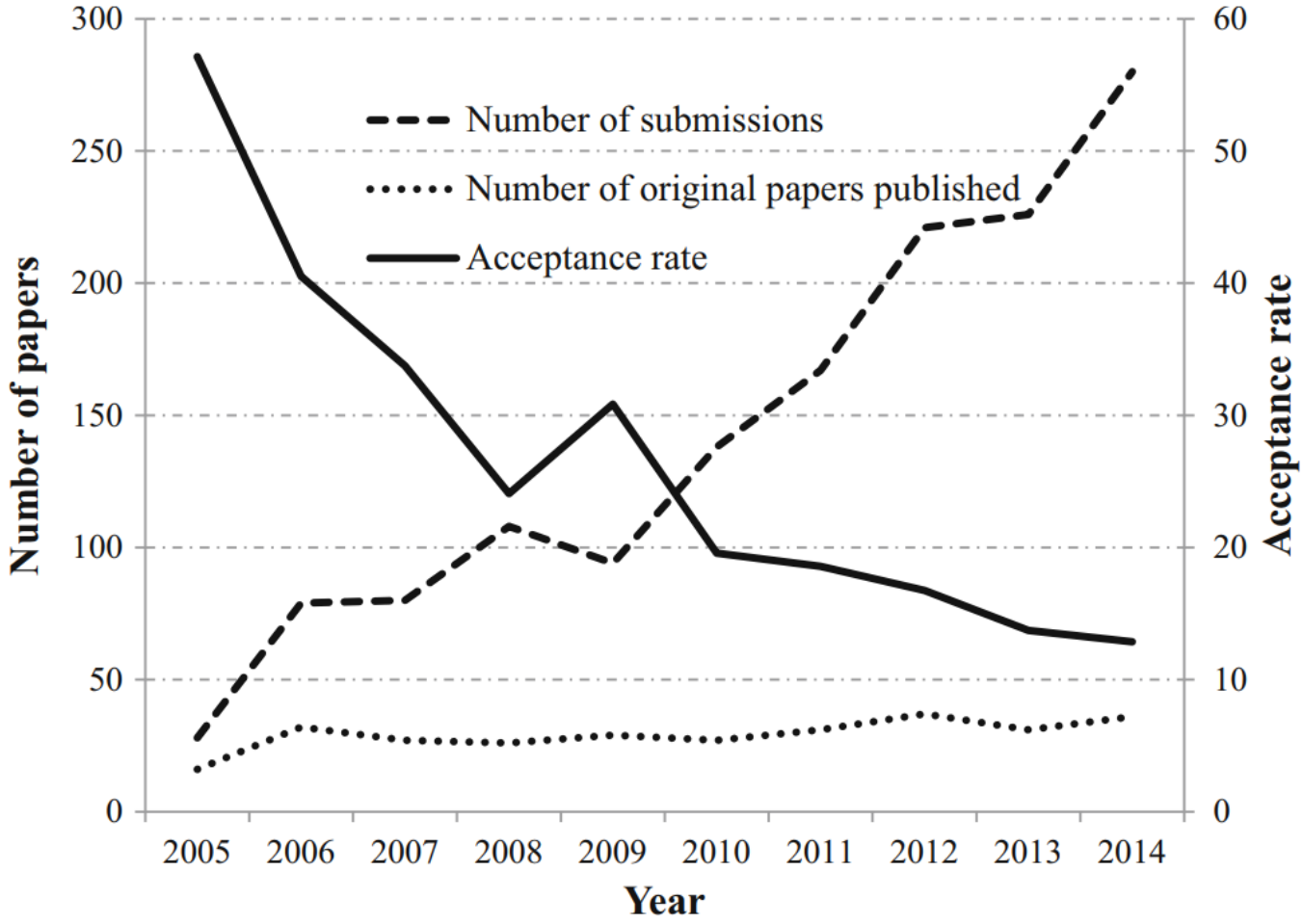
Recent Trends



Experimental papers as a fraction of all papers published in the "top-5" (2015–2018)

Figure: Nikiforakia and Slonim (2019)

Recent Trends



Publication trends in *Experimental Economics* (2005–2014)

Figure: Nikiforakia and Slonim (2015)

Recent Trends

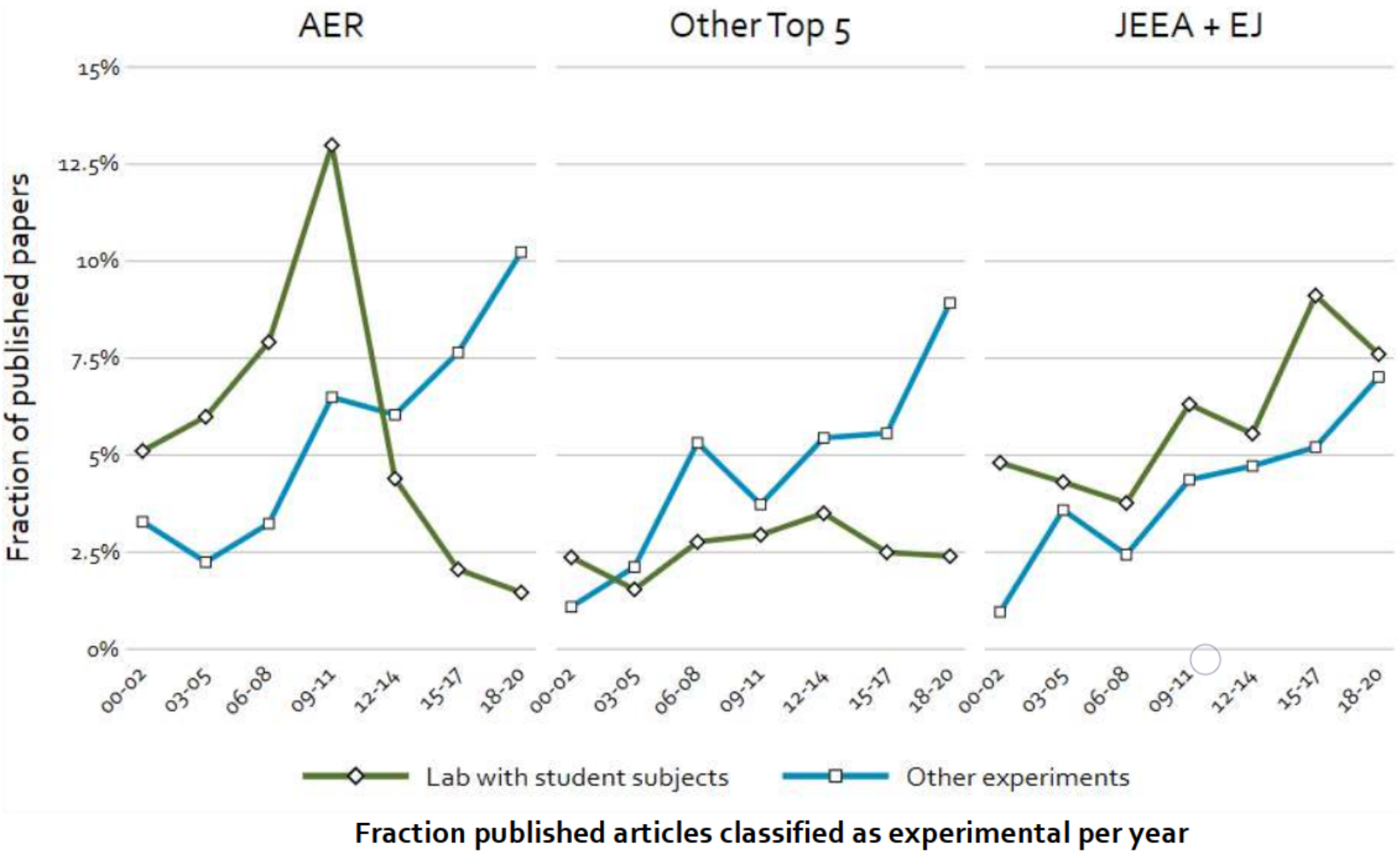


Figure: Reuben et al. (2021)

Recent Trends

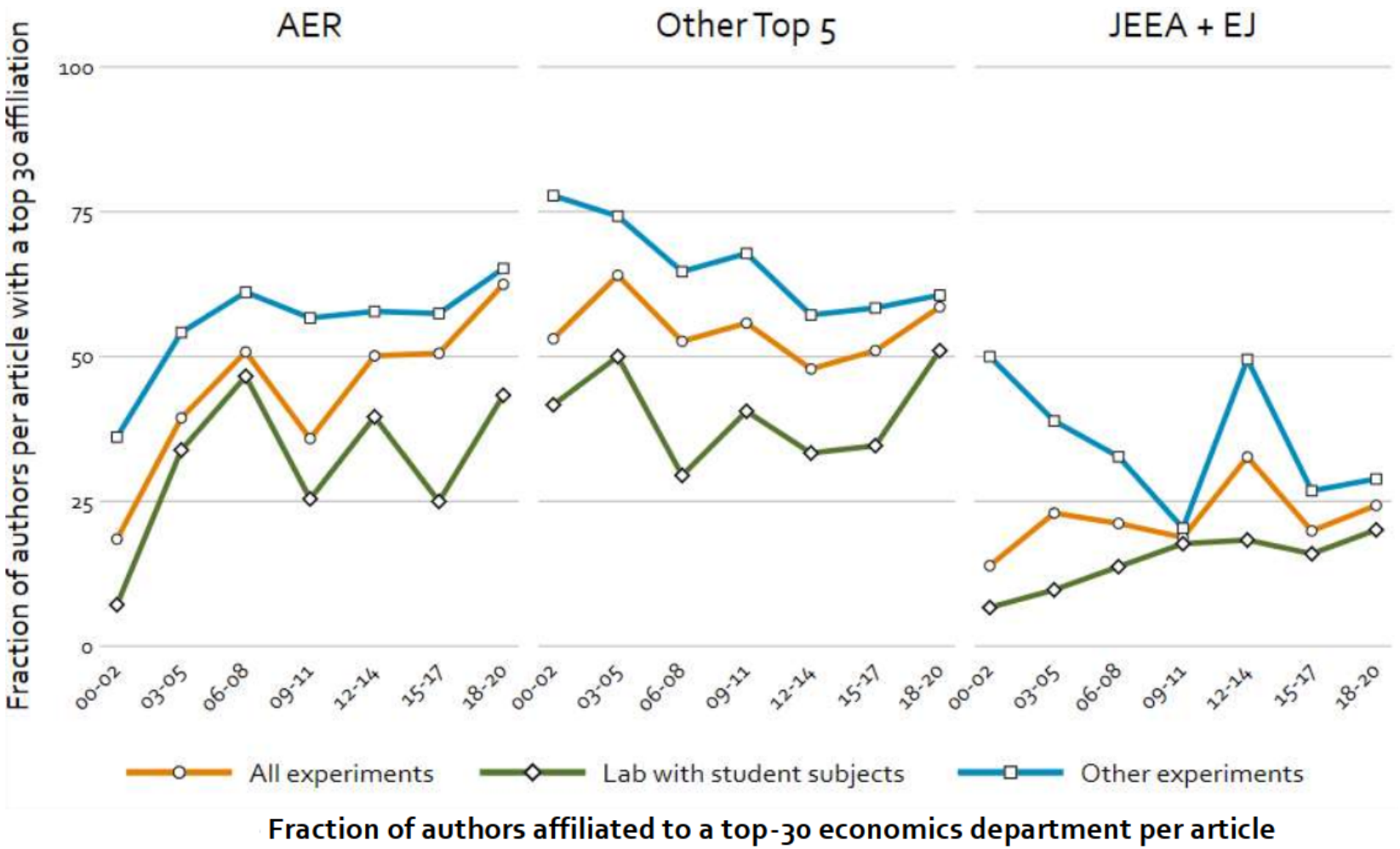


Figure: Reuben et al. (2021)

Recent Trends

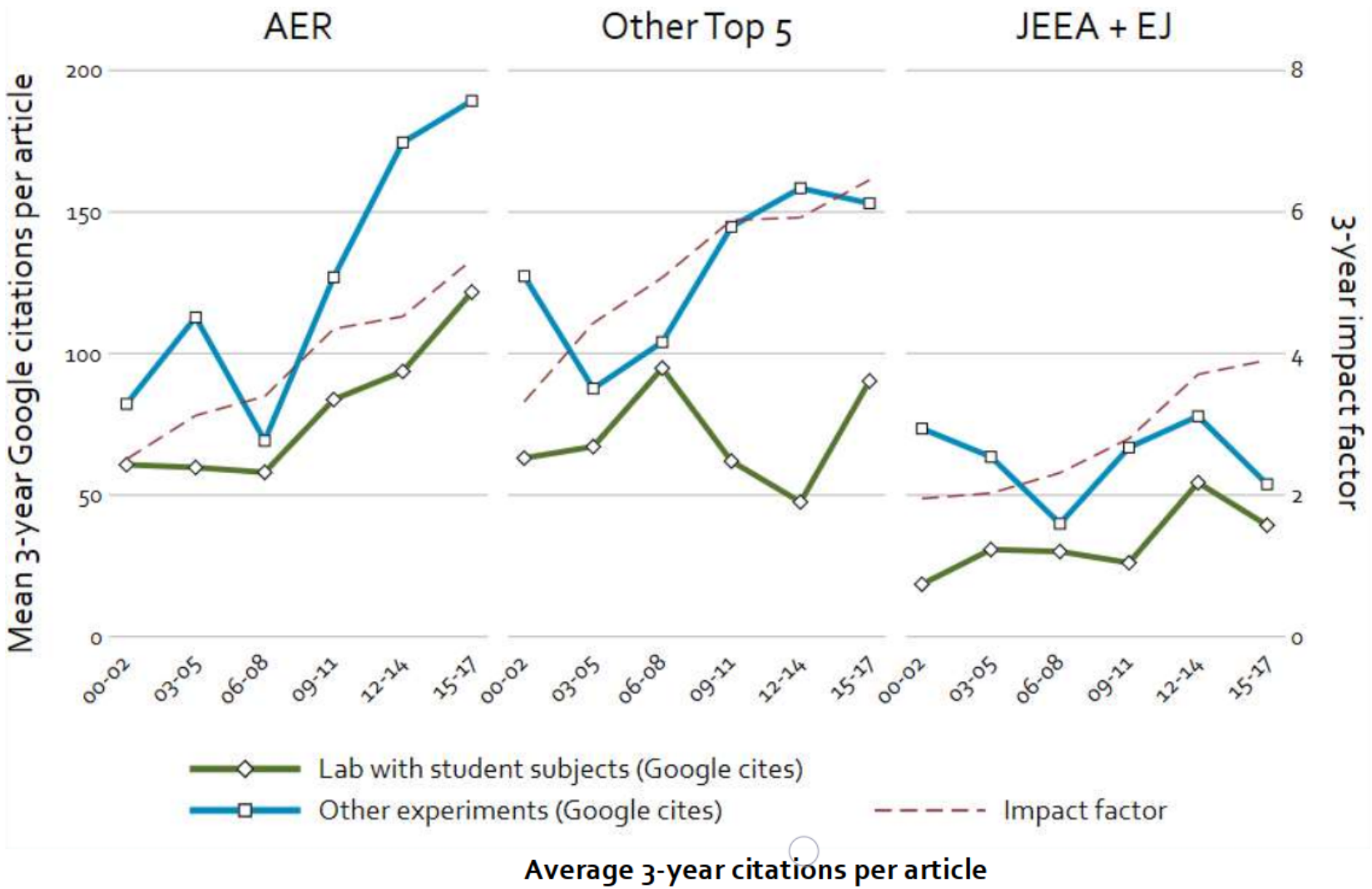


Figure: Reuben et al. (2021)

Methodology Differences

Hertwig and Ortman (2001): economic experiments are superior to psychological experiments in terms of

- 1) Script enactment (with clear instructions);
- 2) Proscription against deception (about the purposes of the experiment);
- 3) Performance-based monetary payments (as incentives);
- 4) Repeated trials (involving learning and feedback).

Why these four differences?

Induced Value Theory

Induced value vs. Homegrown value

- Induced value: experimenter-controlled, exogenously pre-assigned value, publicly known to both the experimenter and participants;

  e.g., Vickrey auction in the lab.
- Homegrown value: unobservable, endogenously subjective value, only known by the participant himself.

  e.g., Vickrey auction in the real-world auction house.

Induced Value Theory

Smith (1976, 1982)

> *"Proper use of a* reward medium *allows an experimenter to induce pre-specified characteristics, and the subjects innate characteristics become largely irrelevant."*

Three sufficient conditions of Induced Value Theory

- 1) Monotonicity or Nonsatiation
- 2) Dominance
- 3) Saliency

Induced Value Theory

1) Monotonicity or Nonsatiation (Smith, 1976, 1982)

- Subjects must prefer higher amounts of the reward medium.
- Subject's preference $V(m, z)$ over reward medium m and other arguments z has $\frac{\partial V(m,z)}{\partial m} > 0$.

Reason for: 2) Proscription against deception (about the purposes of the experiment); 3) Performance-based monetary payments (as incentives).

Induced Value Theory

2) Dominance (Smith, 1976, 1982)

- *"The reward structure dominates any subjective costs (or values) associated with participation in the activities of an experiment."*
- Subjects' utility should be mainly affected by the reward medium m, rather than other factors z, such as fatigue, boredom, experimenter's demand effect, game value, preferences for winning, etc.

Reason for: 2) Proscription against deception (about the purposes of the experiment); 3) Performance-based monetary payments (as incentives).

Induced Value Theory

3) Saliency (Smith, 1976, 1982)

- The reward earned by subjects must depend on his/her choices and possibly the ones of the other subjects as specified by the experiment.
- *"In order that subjects rewards in a laboratory experiment have motivational relevance such rewards must be associated indirectly with the message actions of subjects ... Individuals are guaranteed the right to claim a reward which is increasing (decreasing) in the good (bad) outcomes...of an experiment (no flat payment)."*

Reason for: 1) Script enactment (with clear instructions); 2) Proscription against deception (about the purposes of the experiment); 4) Repeated trials (involving learning and feedback).

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.

Experimental Data

Economic experiments have a growing set of data collection tools.

- 1) Tools developed by social psychologists to measure outcomes, e.g., correspondence study and implicit association test (IAT) to measure discrimination and prejudice;
- 2) Tools inspired by economic theory, e.g., Becker-DeGroot-Marshak (BDM) mechanism to elicit willingness to pay;
- 3) Biomarkers in health, e.g., cortisol to measure stress;
- 4) Eye-tracking method to measure attention;
- 5) Wearable or portable devices to measure mobility or effort, etc.

Data Attributes

Randomized experiments

- Controlled data.

Observational studies

- ? data.

Data Attributes

Randomized experiments

- Controlled data.

Observational studies

- Naturally-occurring data.

Controlled Data			Naturally-Occurring Data	
Lab	AFE	FFE	NFE	NE, PSM, IV, STR

- Lab: Lab experiment
- AFE: Artefactual field experiment
- FFE: Framed field experiment
- NFE: Natural field experiment
- NE: Natural experiment
- PSM: Propensity score estimation
- IV: Instrumental variables estimation
- STR: Structural modeling

Assignment Mechanisms

Randomized experiments

- The assignment mechanism does not depend on observed or unobserved characteristics of the units (i.e., subjects in the experiment);
- The researcher has control over the assignments.

Observational studies

- The assignment mechanism ? depend on observed and/or unobserved characteristics of the units in the study;
- The researcher ? have control over the assignment mechanism.

Assignment Mechanisms

Randomized experiments

- The assignment mechanism does not depend on observed or unobserved characteristics of the units (i.e., subjects in the experiment);
- The researcher has control over the assignments.

Observational studies

- The assignment mechanism may depend on observed and/or unobserved characteristics of the units in the study;
- The researcher does not have control over the assignment mechanism.

Empirical Methods for Estimating Causal Effects

Randomized experiments

- Controlled laboratory experiments and field experiments.

Observational studies

- ?.

Empirical Methods for Estimating Causal Effects

Randomized experiments

- Controlled laboratory experiments and field experiments.

Observational studies

- Instrumental variables (IV), difference-in-differences (DID), regression discontinuity designs (RDD), propensity score matching (PSM), synthesis control, etc.

Statistical Inferences

Randomized experiments

- The **randomization-based approach** takes the subject's potential outcomes as fixed, and considers the assignment of subjects to treatments as random.

$$Y(T) = \begin{cases} Y(1) & \text{if } T = 1 \\ Y(0) & \text{if } T = 0 \end{cases}$$

Observational studies

- The **sampling-based approach** considers the treatment assignments to be fixed, while the outcomes are random;
- Inference is based on the idea that the subjects are a random sample from a much larger population.

$$Y(X) = X\beta + \epsilon$$

Table of Contents

Subject Populations

Frechette (2016): the subject populations can influence our research motivation and experiment design.

- 1) Convenient sample (most common): e.g., undergraduate students;
- 2) Representative sample: highly demographically varied sample representing the larger population, e.g., MTurk online workers;
- 3) Uncommon sample: e.g., children, the elderly, and professionals;
- 4) Infrahuman sample: e.g., rats, pigeons, and monkeys.

Convenient Sample

- The sample is taken from the part of population easy to contact or to reach;
- Convenience sampling does not require a random sample to be generated, since the only criterion is whether the participants agree to participate.

Convenient Sample

Methodological issues:

- The results of the convenience sampling may not be generalized to the target population due to under-representation of the population of interest;

  e.g., undergraduates students might have different behaviors from professional bidders in auction games.

- Homogeneous convenience sampling has insufficient power to identify differences of population subgroups.

  e.g., undergraduates students have similar education and ages.

Professionals

- There is a tendency to assume market forces should lead professionals to be unbiased optimizers and a sample of professionals should confirm standard models;
- A sample of professionals could be used to test predictions of theories and models;

  e.g., observe the behaviors of professional traders in financial markets, test auction theories with professional bidders, play the centipede game with chess players.
- However, professionals does not mean they are necessarily highly sophisticated or unbiased.

Professionals

Methodological issues:

- It is more difficult to recruit a representative sample of professionals, and the costs will be high to provide appropriate incentives to professionals;
- The replicability of experiments with professionals is severely reduced;
- Professionals might behave as if they are in their professional environment, even when this is not appropriate in some experiments.

Children

- Children have no (or few) market interactions, some early socialization, and possibly a still developing neurological structure;
- If a behavior is observed in young children as well as in adults, then it is a **robust** phenomena, as children have less exposure to culture and market institutions;
- If a behavior is not observed in young children but in older children and adults, then it suggests a **learned** behavior;
- Studying both children and their parents can help understand the process of cultural transmission and the origins of economic preferences (i.e., learned from parents or transmitted genetically).

Children

Methodological issues:

- It is difficult to incentivize children who have no market experience. Money might be not ideal. One method is to pay children with tokens that can be used to buy snacks or toys, and show them before doing the experiment;
- It is impractical to explain complicated tasks and certain methods to children clearly, e.g., Becker-DeGroot-Marschak (BDM) mechanism to elicit willingness to pay.

The Elderly

- The elderly have market interactions, the same socialization as the adults, and possibly decaying neurological structure;
- It allows us to gain an understanding of specific brain functions on behavior, since aging does not affect all brain functions equally;
- If a behavior is observed in the standard subject pool and in the elderly, this is evidence of a robust phenomenon;
- Understanding how preferences are shaped by ageing can help design optimal public policies.

The Elderly

Methodological issues:

- It is often inconvenient to bring the elderly to a laboratory. Many of the experiments with the elderly are run online or through the mail;
- Interacting via computers or internet might be difficult for many elderly. Sometimes, abstract instructions can be difficult for them to follow.

Table of Contents

Internal Validity

Internal validity is the ability of a study to estimate **causal** effects within the study population.

- A study has internal validity if the observed correlation between a treatment and an outcome reflects a causal relationship in which the variables are manipulated;
- The internal validity problem is easily solved in controlled experiments because they force the manipulation of treatments to come before the measurement of outcomes.

In general, for internal validity

- Controlled lab experiments ? field experiments ? observational studies

Internal Validity

Internal validity is the ability of a study to estimate **causal** effects within the study population.

- A study has internal validity if the observed correlation between a treatment and an outcome reflects a causal relationship in which the variables are manipulated;
- The internal validity problem is easily solved in controlled experiments because they force the manipulation of treatments to come before the measurement of outcomes.

In general, for internal validity

- Controlled lab experiments $>=$ field experiments $>$ observational studies

External Validity

External validity is concerned with **generalizing** causal inferences drawn for a particular population and setting to alternative settings.

- These alternative settings could involve different populations, different outcomes, or different contexts;
- External validity cannot be guaranteed, neither in randomized experiments, nor in observational studies.

In general, for external validity

- Controlled lab experiments ? field experiments ? observational studies

External Validity

External validity is concerned with **generalizing** causal inferences drawn for a particular population and setting to alternative settings.

- These alternative settings could involve different populations, different outcomes, or different contexts;
- External validity cannot be guaranteed, neither in randomized experiments, nor in observational studies.

In general, for external validity

- Controlled lab experiments $<$ field experiments $<$ observational studies

Deception

No deception in economic experiments:

- Davis and Holt (1993): *"The researcher should ... be careful to avoid deceiving participants. Most economists are very concerned about developing and maintaining a reputation among the student population for honesty in order to ensure that subject actions are motivated by the induced monetary rewards rather than by psychological reactions to suspected manipulation. Subjects may suspect deception if it is present. Moreover, even if subjects fail to detect deception within a session, it may jeopardize future experiments if the subjects ever find out that they were deceived and report this information to their friends."*

Deception

Examples of deceptions:

- Tell subjects that they are paid based on performance and then give them a flat payment;
- Tell subjects that they are randomly matched when they are not;
- Tell subjects that they are matched with 100 other participants when they are matched with 5;
- Tell subjects that they are interacting with a human opponent when they are interacting with a computerized one; etc.

Hertwig and Ortmann (2008): *"... intentional provision of misinformation is deception and that withholding information about research hypotheses, the range of experimental manipulations, or the like ought not to count as deception."*

Noncompliance

Subjects take **treatment different from the assigned one** (e.g., in a study of the effect of vitamin A supplements on infant mortality in rural China, subjects in the control group takes vitamin A supplements).

- 1) Never-takers

  Subjects who would not receive the treatment regardless of the assignment;
- 2) Compliers

  Subjects who would receive the treatment that they are assigned to;
- 3) Always-takers

  Subjects who would receive the treatment regardless of the assignment;
- 4) Defiers

  Subjects who receive the opposite treatment she is assigned.

Noncompliance

- Noncompliance behavior is self-selected and thus ruins the initial randomization;
- Noncompliance reduces power by reducing the contrast between treatments and control;
- Unlike "always-takers" and "never-takers", which just reduce power, "defiers" can bias the experimental results.

Attrition

Attrition is **the loss of participants** in the later phases of a longitudinal study (e.g., in a study of the effect of vitamin A supplements on infant mortality in rural China, some subjects cannot be contacted in post-experiment stage).

- Attrition includes dropout, nonresponse, withdrawal, etc.;
- Non-random attrition might cause attrition bias.

Reactivity

The **observer-expectancy effect** occurs when a researcher's cognitive bias causes them to subconsciously influence the participants of an experiment.

- It may include conscious or unconscious influences on subject behavior including creation of demand characteristics that influence subjects, and altered or selective recording of experimental results themselves.
- **Confirmation bias** can lead to the experimenter interpreting results incorrectly because of the tendency to look for information that conforms to their hypothesis, and overlook information that argues against it.
- It is a significant threat to a study's internal validity, and is therefore typically controlled using a double-blind experimental design.

Reactivity

The **subject-expectancy effect** occurs when a subject expects a given result and therefore unconsciously affects the outcome, or reports the expected result.

- The subject-expectancy effect is most commonly found in medical research, where it can result in the subject experiencing the **placebo effect**.
- Because this effect can significantly bias the results of experiments (especially on human subjects), single-blind or double-blind methodology is used to eliminate the effect.

Reactivity

In a blind experiment, information which may influence the participants of the experiment is withheld until after the experiment is complete.

- Single-blind studies blind participants to their treatment allocation;
- Double-blind studies blind both participants and researchers to treatment allocations.

Reactivity

The **Hawthorne effect** occurs when individuals modify an aspect of their behavior in response to their awareness of **being observed**.

- The original research involved workers who made electrical relays at the Hawthorne Works, a Western Electric plant.
- In the famous lighting study, worker productivity increased with almost any change in the lighting.
- The increased attention from being observed could lead to temporary increases in workers' productivity.

Duhem–Quine Problem

The standard experimental paper uses the following format (Smith, 2002):

- 1) State the theory (T);
- 2) Implement it in a particular context;
- 3) Summarize the implications in testable hypotheses (H);
- 4) Describe the experimental design;
- 5) Present the data and results of the hypothesis tests;
- 6) Conclude that the experiments reject or fail to reject the theoretical hypothesis.

Duhem–Quine Problem

Duhem-Quine (D-Q) problem: it is impossible to test a scientific hypothesis in isolation, because an empirical test of the hypothesis requires one or more **auxiliary hypotheses** (i.e., background assumptions).

- Experimental results always present a joint test of the theory that motivated the test, and all the things you had to do to implement the test (Smith, 2002).
- Thus, if theoretical hypothesis H is implemented with context specific auxiliary hypotheses required to make the test operational, $A_1, A_2, ..., A_n$; then it is $(H|A_1, A_2, ..., A_n)$ that implies observation C.
- If you observe not-C, this can be because any of $(H; A_1, ..., A_n)$ can represent what is falsified.

Therefore, the interpretation of observations in relation to a theoretical hypothesis is inherently ambiguous.

Duhem–Quine Problem

Smith (2002) suggests:

- If statistical test outcomes support not-H (the hypothesis), we tend to conclude that the theory T is "falsified."
- *"But this is not what we do; rather we ask if there is a flaw in the test, i.e., not-A is supported, and we do more experiments. This is good practice — much better than falsificationism."*

When Theory Fails the Tests

When the economic theory fails our experimental tests, what auxiliary hypotheses might we question?

- e.g., the two-person game theory has not done well in tests of ultimatum, dictator, and trust games.

Smith (2010): any combination of **the assumptions** of the theory and its implementation in the experiment may be problematic.

When Theory Fails the Tests

- 1) Backward induction

  Subjects will use backward induction to analyze their situation, to determine the consequences of one's own and the choices of others, and to choose accordingly.

- 2) Independence of history and future

  Subjects make choices in a stage game played once based on what we understand by it — a single play through a decision tree between "strangers" that is devoid of a history and future.

- 3) Complete information

  Subjects have complete information on own and other payoff.

- 4) Domination

  Given a choice between any two amounts of money, subjects always choose the larger amount whatever the circumstances.

When Theory Fails the Tests

- 5) Folk theorem

  Repeat play of a stage game favors cooperation because in "long" games subjects can rationally use punishment to induce domination opponents to play cooperatively.
- 6) Context irrelevance

  Context does not matter, only the underlying abstract game structure.
- 7) Equivalence of the standard game forms

  The two standard game forms (extensive and normal) are equivalent.
- 8) Own vs. other people's money

  Monetary payoffs matter, but not who provides the money or how people acquired the stakes — there is no OPM (other people's money) problem.

Replication

Replication crisis: many scientific studies are difficult or impossible to replicate or reproduce.

- Camerer et al. (2016)

  One-third of 18 experimental studies from two top-tier economics journals (*AER* and *QJE*) failed to replicate.

- Ioannidis et al. (2017)

  "The majority of the average effects in the empirical economics literature are exaggerated by a factor of at least 2 and at least one-third are exaggerated by a factor of 4 or more."

Replication

Replication results.

Plotted are 95% CIs of replication effect sizes (standardized to correlation coefficients). The standardized effect sizes are normalized so that 1 equals the original effect size (fig. S1 shows a nonnormalized version). Eleven replications have a significant effect in the same direction as in the original study [61.1%; 95% CI = (36.2%, 86.1%)]. The 95% CI of the replication effect size includes the original effect size for 12 replications [66.7%; 95% CI = (42.5%, 90.8%)]; if we also include the study in which the entire 95% CI exceeds the original effect size, this increases to 13 replications [72.2%; 95% CI = (49.3%, 95.1%)]. AER denotes the *American Economic Review* and QJE denotes the *Quarterly Journal of Economics*.

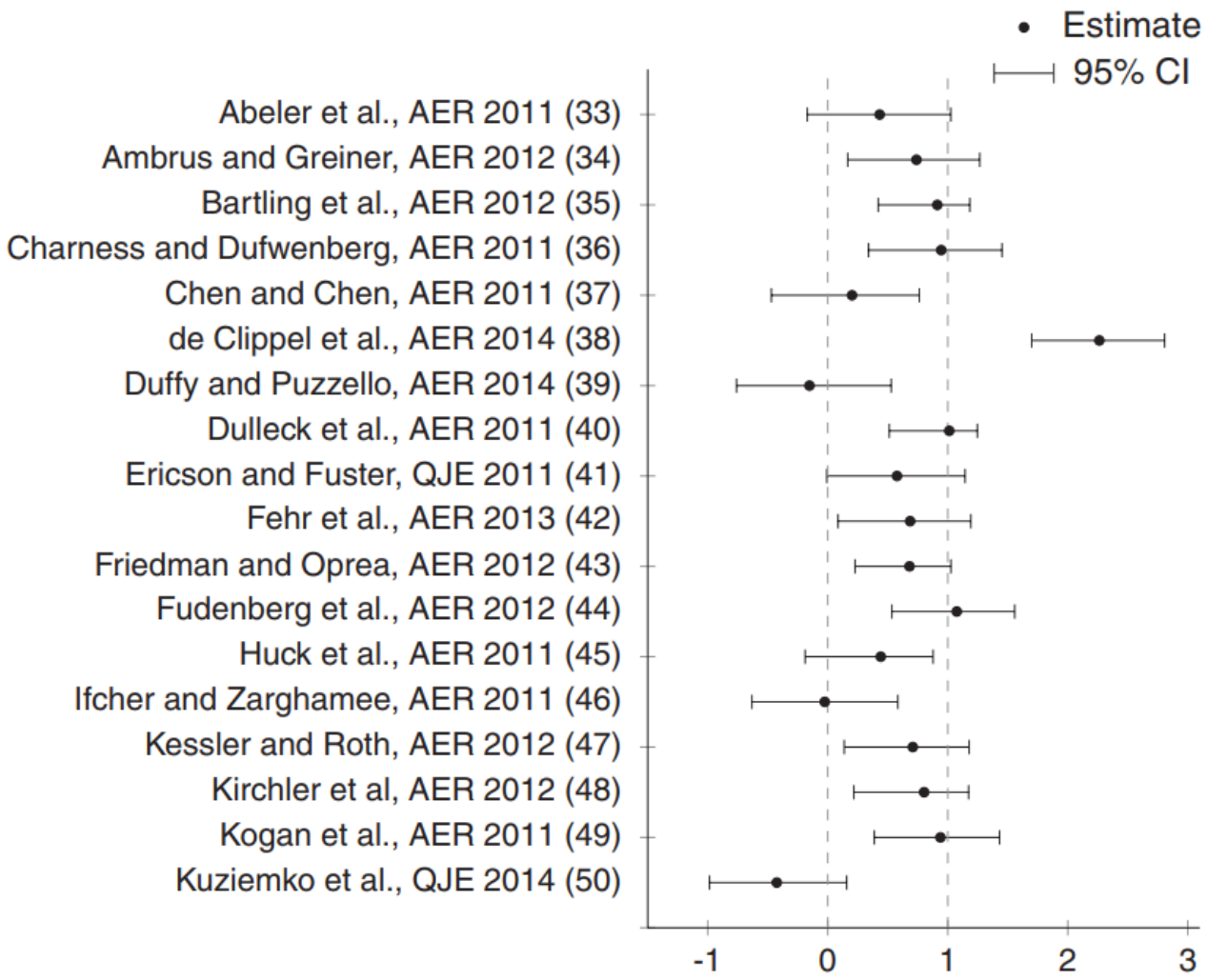


Figure: Camerer et al. (2016)

Replication

Three levels of experiment replication

- 1) Take the actual data generated by an experiment and reanalyze the data to confirm the original findings;
- 2) Run an experiment which follows a similar protocol to the first experiment to determine whether similar results can be generated using new subjects;
- 3) Test the hypotheses of the original study using a new research design.

The 2nd type of replication (i.e., rerun the original experiment, but on a new sample of subjects) can be more difficult.

Replication

Best practices in replication (Chen et al., 2020)

- 1) Understanding theory;
- 2) Choosing culturally appropriate subject pools;
  - In what country/region was the original study conducted?
  - In what language was the original study conducted?
  - Where was the original study conducted? (e.g., lab, field, online)
  - Who were the participants in the original study (e.g., students, MTurk, representative)?
  - Is my sample culturally appropriate given the experimental instructions?
- 3) Calculating sample size with sufficient statistical power;
  - What is the effect size I am trying to replicate?
  - What is the confidence interval of the original effect?
  - What is the sample size of the original study?
  - What is the sample size and power of my replication study given the original effect size?
  - Is my replication sufficiently powered?
  - Does my power calculation use the lower bound of the confidence interval?

Replication

Best practices in replication (Chen et al., 2020)

- 4) Randomizing across treatment sessions;
  - How do I plan to randomize sessions across treatments?
  - Does my randomization compromise theoretical assumptions?
  - What is the size of an original experimental session?
  - What is the size of my experimental session?
  - What differences between the original study and my study might be expected to influence the size and/or direction of the effect?
  - What steps do I plan to take to test whether the differences listed above will influence the outcome of my replication attempt?
- 5) Obtaining original experimenters' endorsement;
- 6) Evaluating replication.

Ethics

Three principles of experiment ethics

- 1) Respect for persons

  People should be treated as autonomous agents. They have their own goals and have the right and ability to decide the best way to pursue them;

- 2) Beneficence

  Researchers should avoid knowingly doing harm and seek to maximize the benefits and minimize the risks to subjects from research;

- 3) Justice

  Researchers should avoid a situation where one group of people bears the risks associated with research while another group receives the benefits.

Ethics

Respect-for-human-subjects requirements

- 1) Informed consent

  Researchers should explain any risks of harm associated with participating in the study to those involved and gain their consent before proceeding;

- 2) Waiving informed consent

  Informed consent can be waived when the risks to the subject are low and the costs of collecting informed consent are high;

- 3) Protecting confidentiality of information

  As part of informed consent, the subject is usually told that any information they provide will be kept confidential.

Table of Contents

Your First Experiment

- Half of you are randomly chosen to get a windfall gift. You are referred as the **"owners;"**
- For the other half of you who are not given a gift, you are referred as the **"non-owners;"**
- As the owners/non-owners are randomly chosen, there may exist welfare gains from trade;
- To test this, the experimenter would like to elicit from each owner the minimum price she is willing to accept to sell the gift;
- From each non-owner, the experimenter would like to elicit the maximum price she is willing to pay in order to buy the gift.

Becker-DeGroot-Marschak (BDM) Mechanism

- 1) Each owner will write down a price to sell the gift, and each non-owner will write down a price to buy it;

Owner's WTA	Non–Owner's WTP	Owner's WTA	Non–Owner's WTP

Becker-DeGroot-Marschak (BDM) Mechanism

- 2) Next a **random price** between CNY 1 and CNY 10 will be determined by throwing a 10-sided die (i.e., random number generator);
- 3) Each owner-non-owner pair will then transact **at the random price if and only if** the random price is higher than (or equal to) the owner's ask and lower than (or equal to) the non-owner's offer.

Becker-DeGroot-Marschak (BDM) Mechanism

Becker et al. (1964): in BDM mechanism, the utility maximizers are incentivized to **truthfully** reveal their reservation price.

- To state one's own reservation price (WTP for the non-owner and WTA for the owner) is a **dominant strategy** for a utility maximizer.
- If you overstate or understate the reservation price, you will miss some opportunity of desirable transactions or be forced into some undesirable transactions.
- From the non-owner's perspective, the method is equivalent to a Vickrey auction against an unknown bidder.

Endowment Effect

Thaler (1980): people often demand much more to give up an item than they would be willing to pay to acquire it. An item's value increases once it becomes part of an individual's **endowment**.

- Knetsch (1989): experiments using mugs and chocolate bars, 89% endowed with mug chose to keep it, 90% endowed with chocolate bar chose to keep it;
- Kahneman, Knetsch and Thaler (1990): experiments using mugs, 11 should have traded, but only 3 did;
- List (2003): market experience tends to eliminate endowment effect.

Fairness Perception

- Q1. A shortage has developed for a popular model of automobile, and customers must now wait two months for delivery. A dealer has been selling these cars at list price. Now the dealer prices this model at CNY 10,000 **above** list price. Is it acceptable or unfair?
- N = ?; "Acceptable" = ?%; "Unfair" = ?%.

Fairness Perception

- Q2. A shortage has developed for a popular model of automobile, and customers must now wait two months for delivery. A dealer has been selling these cars at a discount of CNY 10,000 **below** list price. Now the dealer sells this model only at list price. Is it acceptable or unfair?
- N = ?; "Acceptable" = ?%; "Unfair" = ?%.

Framing Effect

Framing effect is a cognitive bias where people decide on options based on whether the options are presented as a loss or as a gain.

- People tend to avoid risks when a positive frame is presented but seek risks when a negative frame is presented;
- Gain and loss are defined in the scenario as descriptions of outcomes (e.g., lives lost or saved, money paid or saved, etc.);
- Prospect theory shows that a loss is more significant than the equivalent gain;
- A sure gain is favored over a probabilistic gain, and a probabilistic loss is preferred to a definite loss.

Contingent Valuation Study

Contingent valuation study is a survey-based technique for the valuation of non-market resources by asking people to directly report their WTP to obtain a specified good, or WTA to give up a good, rather than inferring them from observed behaviors in the market.

- Induced value: experimenter-controlled, pre-assigned value, publicly known to both the experimenter and participants;
- Homegrown value: unobservable, subjective value, only known by the participant himself.

Contingent Valuation Study

- Q1. How much would you pay to **eliminate** some risk that gives you a .001 chance of sudden death over the next five years?
- N = ?; Mean = ?.

Contingent Valuation Study

- Q2. How much would you have to be paid to **accept** an additional .001 chance of sudden death over the next five years?
- N = ?; Mean = ?.

There is often a big gap between reported WTP and WTA in hypothetical questionnaires, which is one of the reasons to conduct incentive-compatible experiments using real monetary payoffs instead of surveys.

References I


Susan Athey and Guido W Imbens. "The econometrics of randomized experiments". In: *Handbook of Economic Field Experiments*. Vol. 1. Elsevier, 2017, pp. 73–140.


Abhijit Vinayak Banerjee and Esther Duflo. "An introduction to the "Handbook of Field Experiments."". In: *Handbook of Economic Field Experiments*. Vol. 1. Elsevier, 2017, pp. 1–24.


Colin F Camerer et al. "Evaluating replicability of laboratory experiments in economics". In: *Science* 351.6280 (2016), pp. 1433–1436.


Roy Chen, Yan Chen, and Yohanes E Riyanto. "Best practices in replication: a case study of common information in coordination games". In: *Experimental Economics* 24 (2021), pp. 2–30.

References II


Guillaume R Fréchette. "Experimental economics across subject populations". In: *The handbook of experimental economics* 2 (2016), pp. 435–480.


Rachel Glennerster. "The practicalities of running randomized evaluations: partnerships, measurement, ethics, and transparency". In: *Handbook of Economic Field Experiments*. Vol. 1. Elsevier, 2017, pp. 175–243.


Brit Grosskopf. *Experimental/Behavioural Economics Nuremberg 2018*. 2018.


Steven D Levitt and John A List. "Field experiments in economics: The past, the present, and the future". In: *European Economic Review* 53.1 (2009), pp. 1–18.

References III

- John A List. "Field Experiments: A Bridge Between Lab and Naturally-occurring Data". In: (2007).
- Nikos Nikiforakis and Robert Slonim. *Editors' preface: statistics, replications and null results*. 2015.
- Nikos Nikiforakis and Robert Slonim. *Editors' preface: trends in experimental economics (1975–2018)*. 2019.
- Ryan Oprea and Sevgi Yuksel. *Testing Theories Markets & Emergence*. 2020.
- Daniela Puzzello. *Why and How to Experiment in Economics*. 2019.
- Ernesto Reuben et al. "Trends in the Publication of Experimental Economics Articles". In: *Available at SSRN 3907528* (2021).

References IV

Vernon L Smith. "Method in experiment: Rhetoric and reality". In: *Experimental economics* 5.2 (2002), pp. 91–110.

Vernon L Smith. "Theory and experiment: What are the questions?" In: *Journal of Economic Behavior & Organization* 73.1 (2010), pp. 3–15.