Behavioural Economics

Topics in Experimental Economics

Taha Movahedi

University of Warwick

E-mail: taha.movahedi@*warwick.ac.uk*

Topic 1: About Experimental Economics

Topic 1: About Experimental Economics

1. Brief History of Experimental Economics
2. When to use Experiments
3. A Good Experiment and Experimental Designs
4. Good Practices
5. Internal and External Validity

Experiments

WARWICK

- Experimental methods are standard across most sciences.
- The approach stressed by Karl Popper suggest that “good science” be falsifiable.
- This is the norm in Physics, Chemistry, Biology, etc. but for many years Economics relied on existing datasets without directly testing theory.

Brief History of Experimental Economics

WARWICK

- For a long time experiments were considered impossible:

  "We can seldom test particular predictions in the social sciences by experiments explicitly designed to eliminate what are judged to be the most important disturbing influences" (Friedman, 1953, p. 10).
- Allais (1953)*, Ellsberg (1961)*, Markowitz (1952)*

  anomalous implications of expected and subjective expected utility
- Strotz (1955)*

  following the Expected Utility Theory and the discounted utility models-questioned exponential discounting

Brief History of Experimental Economics

WARWICK

- Kahneman and Tversky (1979)

  developed theory for simple lotteries and stated probabilities that expected utility theory failed to support
- Thaler (1981)* and Loewenstain and Prelec (1992)*

  seminal work on dynamic inconsistency and discounted utility

Brief History of Experimental Economics

WARWICK

- Some of the oldest experiments within economics are older than you might think!
- Market experiments: Decentralized markets - Chamberlain (1948) induced demand and cost structure;

Brief History of Experimental Economics

WARWICK

- Some of the oldest experiments within economics are older than you might think!
- Market experiments: Decentralized markets - Chamberlain (1948) induced demand and cost structure;
  - Aim: replicate the functioning of a market in perfect competition with rational agents
  - Design: Subjects (students) randomly assigned role of either a seller or a buyer
  - Elicited demand (willingness to pay) and supply (willingness to accept)
  - Result: outcomes systematically deviated from competitive predictions

Chamberlain (1948)

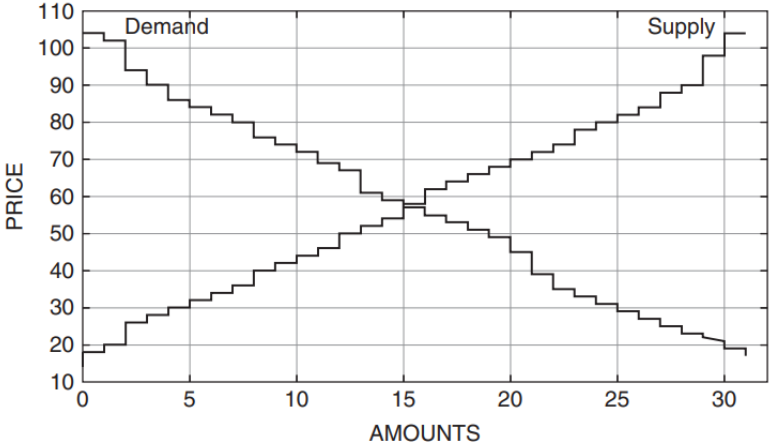


Figure 1.2 Market equilibrium in the Chamberlin (1948) experiment

Note. The figure shows the theoretical equilibrium of the market implemented in the laboratory – at the intersection of the (increasing) supply function and the (decreasing) demand function.
Source: Chamberlin (1948, p. 97, Figure 1).

Figure: Jacquemet, N., & L'Haridon, O. (2018). The Emergence of Experiments in Economics. In

Chamberlain (1948)

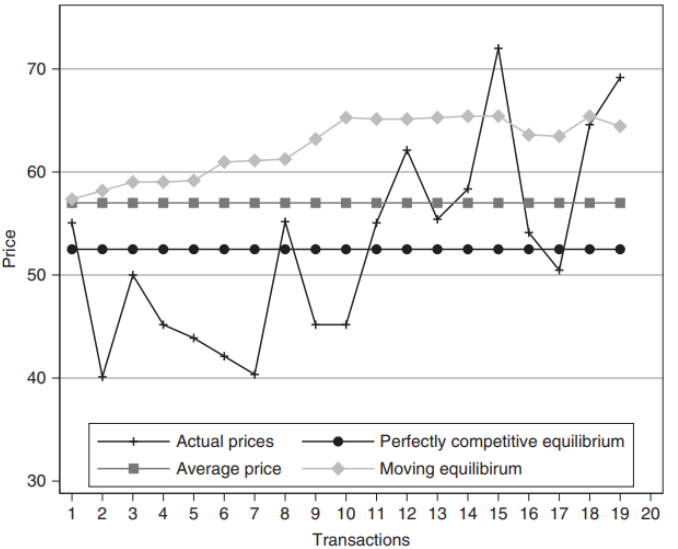


Figure 1.3 Observed behaviour in the Chamberlin (1948) experiment

Note. For each transaction in abscissa, the figure shows the actual price observed in the experiment as well as a recall of the theoretical equilibrium described in Figure 1.2.

Source: Chamberlin (1948, p. 101, Figure 3).

Brief History of Experimental Economics

WARWICK

- Chamberlin to conclude, “Perhaps it is the assumption of a perfect market which is “strange” in the first place” (and interpret this as a support for his monopolistic competition model).
- This result is not, however, the end of the experimental story of markets.

Brief History of Experimental Economics

WARWICK

- Double auction - Vernon Smith (1962, 1964).
  - Smith (1962, 1964) replicated Chamberlin's experiment

Brief History of Experimental Economics

WARWICK

- Double auction - Vernon Smith (1962, 1964).
  - Smith (1962, 1964) replicated Chamberlin's experiment
  - Design: double auction
  - bids, offers, and transactions prices are public information
  - Repeated several rounds
  - Result: markets could converge to efficient, competitive outcomes, even with a small number of traders who initially knew nothing about market conditions

Smith (1962)

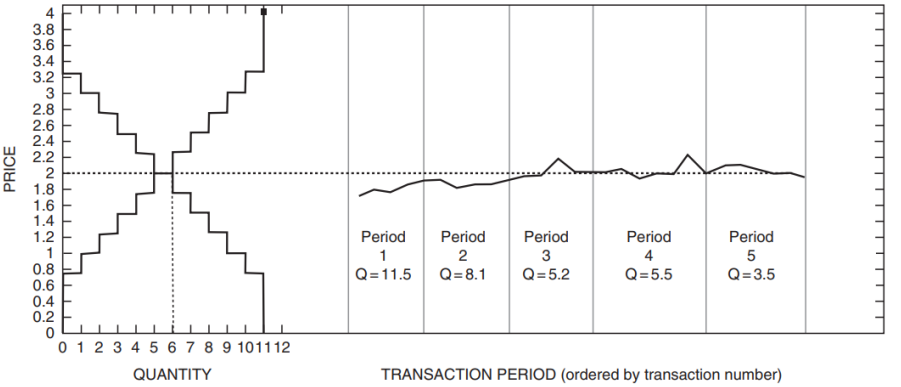


Figure 1.4 Predicted and observed behaviour in the Smith (1962) replication

Note. The left-hand side shows the theoretical market equilibrium – at the intersection of the (increasing) supply function and the (decreasing) demand function. The right-hand side shows the price and number of transactions in each market period.
Source: Smith (1962, p. 113, Figure 1).

Figure: Source: Jacquemet, N., & L'Haridon, O. (2018). The Emergence of Experiments in Economics. In *Experimental Economics: Method and Applications* (pp. 3-25). Cambridge: Cambridge University Press. doi:10.1017/9781107446786.002

Brief History of Experimental Economics

WARWICK

- Game experiments: Prisoners'dilemma 1950's - Originally by psychologists and sociologists
- Oligopoly games Reinhard Selten (1959).

Brief History of Experimental Economics

WARWICK

- A shift in the trend in the first decades of the new millennium
- Nobel Prize in Economics:

Brief History of Experimental Economics

WARWICK

- A shift in the trend in the first decades of the new millennium
- Nobel Prize in Economics:
- In 2002, Vernon L. Smith and Daniel F. Kahneman

  for establishing it as a conclusive method

Brief History of Experimental Economics

WARWICK

- A shift in the trend in the first decades of the new millennium
- Nobel Prize in Economics:
- In 2002, Vernon L. Smith and Daniel F. Kahneman

  for establishing it as a conclusive method
- Alvin Roth and Lloyd Shapley

  for the theory of stable allocations and the practice of market design

Brief History of Experimental Economics

WARWICK

- A shift in the trend in the first decades of the new millennium
- Nobel Prize in Economics:
- In 2002, Vernon L. Smith and Daniel F. Kahneman

  for establishing it as a conclusive method
- Alvin Roth and Lloyd Shapley

  for the theory of stable allocations and the practice of market design
- Richard Thaler was awarded in 2017

  for having incorporated psychologically realistic assumptions into analyses of economic decision-making

Experiments in Economics

WARWICK

- Experiments are now regularly used in Economics, but they are a little different.
- As a rule of thumb many experimentalists feel that if most subjects follow a theory (perhaps >70% if there are is one theory and one alternative) then the theory is not too bad – note how this differs from a "law of physics" which is instantly disproved by a single failed observation.
- The crucial difference is that in Economics models and assumptions about behaviour are often viewed as simplifications – perhaps for tractability and to allow predictions to be made – and not as "true".

Experiments in Economics

WARWICK

- Running experiments is now an established method to explain and/or describe economic activity which bring economics into alignment with many of the natural sciences which rely on experimental methods (e.g. physics and biology).
- From the last 12 years 11% of the most-cited papers are experimental which is roughly the same number as theoretical papers with the big shift towards experiments coming towards the end of the 20th century.

Experiments in Economics

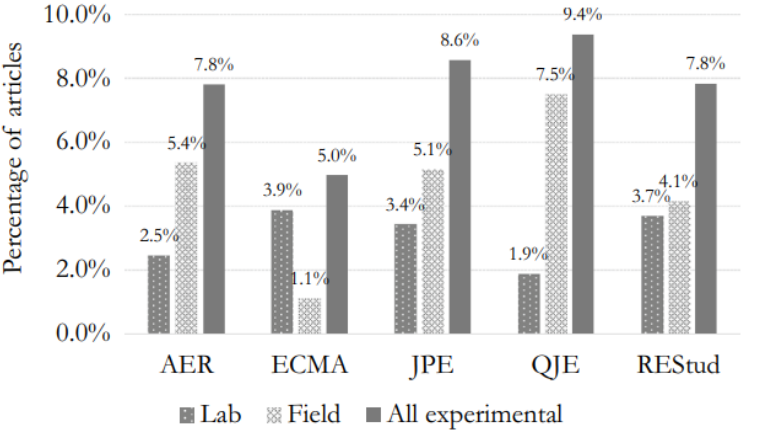


Figure: Source: Experimental papers as a fraction of all papers published in the "top-5" (2015–2018). Source: Nikiforakis & Slonim (2019)

See also Reuben et al. (2021)

Experimental Economics

WARWICK

- What It Is, Why: An Overview
- How: Designing An Experiment, Incentives, Internal-Validity Issues
- What For: External Validity of Experimental Results,
- Final Notes

When to Use Experiments?

- Identify the main driving forces of behaviour.

When to Use Experiments?

- Identify the main driving forces of behaviour.

  Psychological questionnaires

  Non-standard preferences

When to Use Experiments?

- Identify the main driving forces of behaviour.

  Psychological questionnaires

  Non-standard preferences
- Testing theories, Game theory and assessing their empirical success

  providing control over institutions, personal characteristics, incentives, randomization.

When to Use Experiments?

- Identify the main driving forces of behaviour.

  Psychological questionnaires

  Non-standard preferences
- Testing theories, Game theory and assessing their empirical success

  providing control over institutions, personal characteristics, incentives, randomization.
- Looking for empirical regularities, Challenging theory and Evaluating Assumptions: rationality, NE, EUT

When to Use Experiments?

- Identify the main driving forces of behaviour.

  Psychological questionnaires

  Non-standard preferences
- Testing theories, Game theory and assessing their empirical success

  providing control over institutions, personal characteristics, incentives, randomization.
- Looking for empirical regularities, Challenging theory and Evaluating Assumptions: rationality, NE, EUT
- Informing theory (e.g. the role of experiments in developing behavioural science)
- Policy Evaluation (e.g. spectrum auctions)

When to Use Experiments?

- Identify the main driving forces of behaviour.

  Psychological questionnaires

  Non-standard preferences
- Testing theories, Game theory and assessing their empirical success

  providing control over institutions, personal characteristics, incentives, randomization.
- Looking for empirical regularities, Challenging theory and Evaluating Assumptions: rationality, NE, EUT
- Informing theory (e.g. the role of experiments in developing behavioural science)
- Policy Evaluation (e.g. spectrum auctions)
- Elicitation procedures

What for: Testing Theories

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.

What for: Testing Theories

WARWICK

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:

What for: Testing Theories

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:
  - Institutions (e.g. voting rules, communication, etc.) though not perfect (e.g. social norms).

What for: Testing Theories

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:
  - Institutions (e.g. voting rules, communication, etc.) though not perfect (e.g. social norms).
  - Incentives (payoffs) again not perfect (e.g. altruism).

What for: Testing Theories

WARWICK

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:
  - Institutions (e.g. voting rules, communication, etc.) though not perfect (e.g. social norms).
  - Incentives (payoffs) again not perfect (e.g. altruism).
  - Measure and checking confounding or unobservable (in the field) variables (e.g. beliefs).

What for: Testing Theories

WARWICK

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:
  - Institutions (e.g. voting rules, communication, etc.) though not perfect (e.g. social norms).
  - Incentives (payoffs) again not perfect (e.g. altruism).
  - Measure and checking confounding or unobservable (in the field) variables (e.g. beliefs).
  - Randomization (avoids some self-selection problems).

What for: Testing Theories

WARWICK

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics and/or measured conditions that are typically unavailable in field data.
- The key thing that experimental economics provides is control of various forms:
  - Institutions (e.g. voting rules, communication, etc.) though not perfect (e.g. social norms).
  - Incentives (payoffs) again not perfect (e.g. altruism).
  - Measure and checking confounding or unobservable (in the field) variables (e.g. beliefs).
  - Randomization (avoids some self-selection problems).

What for: Game Theory

WARWICK

- The development of Game Theory gave particular impetus to experimental economics in the 1950's, as game theory offered testable theories of economic behaviour that depended on the fine structure of both the strategic environment and the preferences of the players.
- Both of these aspects of a game require very high levels of control and are unlikely to be seen in the "real world".
- It is therefore no surprise that many experimental economists are also (game) theorists often seeking to test their own theories or those of their peers.

What for: Empirical Regularities

WARWICK

- Experimental data can allow us to search for regularities, and exploring and documenting unanticipated regularities has given experimental economics some of it's biggest hits (from Allais onwards).
- These are often all about violations of the predictions of existing theories which explains the links with behavioural economics.

What for: Empirical Regularities

WARWICK

- Experimental data can allow us to search for regularities, and exploring and documenting unanticipated regularities has given experimental economics some of it's biggest hits (from Allais onwards).
- These are often all about violations of the predictions of existing theories which explains the links with behavioural economics.
- Challenging theory: For example, do people play Nash strategies, do they really never play dominated strategies, can they optimize decisions, use Bayes rule, calculate risks properly, solve the sorts of problems implicit in economic theory?

What for: Empirical Regularities

WARWICK

- Experimental data can allow us to search for regularities, and exploring and documenting unanticipated regularities has given experimental economics some of it's biggest hits (from Allais onwards).
- These are often all about violations of the predictions of existing theories which explains the links with behavioural economics.
- Challenging theory: For example, do people play Nash strategies, do they really never play dominated strategies, can they optimize decisions, use Bayes rule, calculate risks properly, solve the sorts of problems implicit in economic theory?
- Helping theory: For example, if there are multiple equilibria which is selected? What assumptions really matter.

What for: Evaluating Assumptions

WARWICK

- Theory often contains assumptions, for instance: as N increases such and such happens.
- We may want to explore the gap between what is necessary for a theory to hold and what is sufficient: for instance when N is infinite firms behave as competitive - but how big a number is really enough?

What for: Evaluating Assumptions

WARWICK

- Theory often contains assumptions, for instance: as N increases such and such happens.
- We may want to explore the gap between what is necessary for a theory to hold and what is sufficient: for instance when N is infinite firms behave as competitive - but how big a number is really enough?
- Where theory is questioned there is also the issue of when and why and this relates to so-called "stress-testing": a theory may fail for a certain set of parameters - will it do better with others? For instance how do contributions to a public good game change as N rises?

What for: Evaluating Assumptions

WARWICK

- Theory often contains assumptions, for instance: as N increases such and such happens.
- We may want to explore the gap between what is necessary for a theory to hold and what is sufficient: for instance when N is infinite firms behave as competitive - but how big a number is really enough?
- Where theory is questioned there is also the issue of when and why and this relates to so-called "stress-testing": a theory may fail for a certain set of parameters - will it do better with others? For instance how do contributions to a public good game change as N rises?
- In this way experimental economics may help to be clearer on when theory works and when it does not.

What for: Informing Theory

WARWICK

- As well as helping to refine theoretical ideas experiments can also assist in the formulation of new theories, to explain newly observed regularities, and devising new experiments to help distinguish among such theories.
- Examples: behavioural economics, generalized and alternative models of expected utility theory, learning in games, bounded rationality, etc.

What for: Informing Theory

WARWICK

- As well as helping to refine theoretical ideas experiments can also assist in the formulation of new theories, to explain newly observed regularities, and devising new experiments to help distinguish among such theories.
- Examples: behavioural economics, generalized and alternative models of expected utility theory, learning in games, bounded rationality, etc.
- Every time an experiment reveals a contradiction with standard theory it might reveal something about behaviour but it could also be because there some flaw in the design.
- Some theorists ignore experimental findings choosing to dismiss experimental economics, but many (most) take it to suggest that there is an empirical basis for behavioural economics.

What for: Policy

- There are also policy-oriented experiments. Most commonly to do with market design.
- There are many examples, but think about something as huge as conducting a spectrum auction or as complex as the effects on worker motivation of various compensation schemes.
- Would it make sense to just go ahead without first testing the mechanism you have invented? (UK spectrum auctions example: Binmore and Klemperer). studies.

What for

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics
- Experiments make a clear conceptual point that can be used by others

What for

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics
- Experiments make a clear conceptual point that can be used by others

- Cleanly identify causal effect of x on y

  Increase in observability increases prosocial giving

What for

- A controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics
- Experiments make a clear conceptual point that can be used by others

- Cleanly identify causal effect of x on y

  Increase in observability increases prosocial giving

- Cleanly measure some parameter (and then work with it)

  People are present-biased

  Measure time preferences and correlate with savings behavior

- Experiments and economic theory go hand in hand: experiments are about assessing the empirical relevance of the behavioural content of economic models.

- Experiments and economic theory go hand in hand: experiments are about assessing the empirical relevance of the behavioural content of economic models.
- Experimental economics and behavioural economics are clearly linked – as suggested experiments have produced many empirical regularities that support behavioural economics.

- Experiments and economic theory go hand in hand: experiments are about assessing the empirical relevance of the behavioural content of economic models.
- Experimental economics and behavioural economics are clearly linked – as suggested experiments have produced many empirical regularities that support behavioural economics.
- However, hopefully you now see that experimental economics is a methodological field (like mathematical economics or econometrics) that can be widely applied and not a part of behavioural economics, though the two are often linked both in teaching and research.

- Experiments and economic theory go hand in hand: experiments are about assessing the empirical relevance of the behavioural content of economic models.
- Experimental economics and behavioural economics are clearly linked – as suggested experiments have produced many empirical regularities that support behavioural economics.
- However, hopefully you now see that experimental economics is a methodological field (like mathematical economics or econometrics) that can be widely applied and not a part of behavioural economics, though the two are often linked both in teaching and research.
- Moreover, experiments can and have also confirmed many conventional (non-behavioural) theories.

A Good Experiment

WARWICK

- Two key ingredients
  - In blatant contradiction to contemporary thinking / theoretical predictions
  - Very simple
- The real goal is a design that offers the best opportunity to learn something useful and to answer the questions that motivate your research.
- An experiment is judged by it's impact on our understanding of behaviour.

A Good Experiment

WARWICK

- Should experiments replicate reality?
- Should an experiment replicate a formal model?

A Good Experiment

WARWICK

- Should experiments replicate reality?
- Should an experiment replicate a formal model?
  - No: the real goal is a design that offers the best opportunity to learn something useful and to answer the questions that motivate your research.
  - An experiment is judged by it's impact on our understanding of behaviour.

A Good Experiment

WARWICK

- A good design:

A Good Experiment

WARWICK

- A good design:

  Is simple compared to reality and even simpler than relevant models (remember that models are themselves a simplification of reality).

A Good Experiment

WARWICK

- A good design:

  Is simple compared to reality and even simpler than relevant models (remember that models are themselves a simplification of reality).

  Is designed to test specific hypothesis or set of hypotheses (note the multiple testing problem and the dangers of spurious correlations).

A Good Experiment

WARWICK

- A good design:

  Is simple compared to reality and even simpler than relevant models (remember that models are themselves a simplification of reality).

  Is designed to test specific hypothesis or set of hypotheses (note the multiple testing problem and the dangers of spurious correlations).

  Tests or controls for alternative hypotheses.

A Good Experiment

WARWICK

- A good design:

  Is simple compared to reality and even simpler than relevant models (remember that models are themselves a simplification of reality).

  Is designed to test specific hypothesis or set of hypotheses (note the multiple testing problem and the dangers of spurious correlations).

  Tests or controls for alternative hypotheses.
- Potential alternative hypotheses may again depend whom you talk to, which is why psychology experiments often look so different to economics experiments.

A Good Experiment: Confounding

WARWICK

- A good design controls for the most plausible alternative hypotheses that might explain what is being observed:
  - This can help us to avoid confounding theories that give an equally plausible rationale for behaviour under the experimental design.
  - And to protect ourselves from fooling ourselves into believing what we want to believe.

A Good Experiment: Confounding

WARWICK

- A good design controls for the most plausible alternative hypotheses that might explain what is being observed:
  - This can help us to avoid confounding theories that give an equally plausible rationale for behaviour under the experimental design.
  - And to protect ourselves from fooling ourselves into believing what we want to believe.
- It might be that experimenters are following up on their intuitions, and will therefore investigate hypotheses that they believe to be true.

A Good Experiment: Confounding

WARWICK

- A good design controls for the most plausible alternative hypotheses that might explain what is being observed:
  - This can help us to avoid confounding theories that give an equally plausible rationale for behaviour under the experimental design.
  - And to protect ourselves from fooling ourselves into believing what we want to believe.
- It might be that experimenters are following up on their intuitions, and will therefore investigate hypotheses that they believe to be true.
- The same intuition that causes you to believe the hypothesis might give you a good idea of situations in which the hypothesis will hold.

A Good Experiment: Confounding

WARWICK

- A good design controls for the most plausible alternative hypotheses that might explain what is being observed:
  - This can help us to avoid confounding theories that give an equally plausible rationale for behaviour under the experimental design.
  - And to protect ourselves from fooling ourselves into believing what we want to believe.
- It might be that experimenters are following up on their intuitions, and will therefore investigate hypotheses that they believe to be true.
- The same intuition that causes you to believe the hypothesis might give you a good idea of situations in which the hypothesis will hold.
- But if there are other reasons that those conclusions might hold, you have to make sure that you haven't just created a situation that gives you the results you expect, but not for the reason that you believe.

A Good Experiment: Confounding

WARWICK

- What are the most plausible alternative hypotheses we should be controlling for?
- There is no recipe for it.
- Read a lot
- Present it a lot
- Run a Pilot first

Designing an Experiment

WARWICK

- The primary strength of the experimental method: control
- Treatments should be compared against a control group

Designing an Experiment

WARWICK

- The primary strength of the experimental method: control
- Treatments should be compared against a control group
- Good practice involves:
  - Testing hypothesis by changing one variable at a time.
  - Only changing variables which are directly relevant to the hypothesis being tested, otherwise holding the environment fixed.
  - Avoiding confounds (don't change more than one thing at a time).

Designing an Experiment

WARWICK

- We can deal with many uncontrolled factors via randomization.

  For example, experiments designed to test how subjects' attitudes towards fairness are affected by some treatment variable.

  Subjects enter the lab with differing attitudes about fairness so a true controlled experiment can't be run but by randomly assigning subjects to treatments, we can eliminate subjects differing attitudes as a cause of differences between treatments.

Designing an Experiment

WARWICK

- We can deal with many uncontrolled factors via randomization.

  For example, experiments designed to test how subjects' attitudes towards fairness are affected by some treatment variable.

  Subjects enter the lab with differing attitudes about fairness so a true controlled experiment can't be run but by randomly assigning subjects to treatments, we can eliminate subjects differing attitudes as a cause of differences between treatments.
- This relies on the law of large numbers, implying that a large sample may be necessary.

Designing an Experiment

WARWICK

- We can deal with many uncontrolled factors via randomization.

  For example, experiments designed to test how subjects' attitudes towards fairness are affected by some treatment variable.

  Subjects enter the lab with differing attitudes about fairness so a true controlled experiment can't be run but by randomly assigning subjects to treatments, we can eliminate subjects differing attitudes as a cause of differences between treatments.
- This relies on the law of large numbers, implying that a large sample may be necessary.
- Or, we can measure variables which you think may affect fairness directly: gender or age for example.

  This explains why collecting demographic information (via a questionnaire) is so standard in experiments.

Within Vs. Between Subject Designs

- Within vs. between (panel vs. cross section) designs allow some indirect experimental control.

Within Vs. Between Subject Designs

WARWICK

- Within vs. between (panel vs. cross section) designs allow some indirect experimental control.
- Within-subject design: participants make decisions in all treatments.
- Between-subject design: different participants make decisions in each treatment.
- Under a within-subject design each subject is its own control.
- However on the flip side there is the disadvantage of order effects or even fatigue.

Matching Designs

WARWICK

- For some game theoretic experiments there is also the issue of how to match up subjects:
- Partners:

Matching Designs

WARWICK

- For some game theoretic experiments there is also the issue of how to match up subjects:
- Partners: always play with same group.
- Strangers:

Matching Designs

- For some game theoretic experiments there is also the issue of how to match up subjects:
- Partners: always play with same group.
- Strangers: randomly re-matched before playing each game.
- Perfect strangers:

Matching Designs

WARWICK

- For some game theoretic experiments there is also the issue of how to match up subjects:
- Partners: always play with same group.
- Strangers: randomly re-matched before playing each game.
- Perfect strangers: subjects do not play with the same subjects more than once.

Multiple Rounds and Repeated Designs

- One round versus many rounds?
- Multiple rounds may be an important part of the theory to be tested.

Multiple Rounds and Repeated Designs

WARWICK

- One round versus many rounds?
- Multiple rounds may be an important part of the theory to be tested.
- but there will be implications for learning.
- One way to overcome this is to train participants, or test them before you use them as participants in your experiment.

  For example, using example rounds or having simple tests of understanding before the “main experiment” begins..

Incentives and Payments

WARWICK

- Experiments in Economics are (almost always) incentivized – we do this because we believe incentives are important (to encourage effort and to incentivize getting decisions right).
- That means we do not just pay a fixed fee for taking part but an amount based on performance.

Incentives and Payments

WARWICK

- So how much to pay?

Incentives and Payments

WARWICK

- So how much to pay?
- Since incentives matter it may be sensible to keep the fixed turnout fee relatively low and make performance-related payments high.
- You may wish to give higher fixed turnout fees if they are an important aspect of the experiment. For instance giving each subject a lump sum to use in the experiment though this means they might leave with nothing.
- Keeping costs down is one reason why experiments in developing countries are so popular since they can help to keep costs down while still providing high incentives relative to standards of living.

Pay Once or Many Times?

WARWICK

- A second payment decision is whether to pay one (randomized) round or all rounds?

Pay Once or Many Times?

- A second payment decision is whether to pay one (randomized) round or all rounds?
- It is now fairly common to inform subjects taking part in a multi-round experiment that one round will be selected at random and it is that and only that round that will have payoff consequences.
- Since any round might be selected this still leaves all the rounds incentivized.
- There is some evidence that even though this perhaps should not matter it can matter – possibly some behavioural biases will come into play.
- Nonetheless this has become a well-used alternative method – you might need to judge on an experiment by experiment basis whether it makes good sense or not.

Neutral Language

Neutral Language

WARWICK

- Use language that is neutral, or frame/prime?
- Example: if you want one session to be “happier” to check for the importance of happiness on decision-making then you can use a “mood- induction procedure” such as showing a video or handing out sweets to “prime” subjects but this is controlled and measurable and will be a variable in the final regression (as a treatment).
- Inadvertent priming or framing can ruin an experimental design.

About the Subject Choices

- Do you want a homogeneous subject group, or would variety be better?

About the Subject Choices

WARWICK

- Do you want a homogeneous subject group, or would variety be better?
- In some cases you are looking for the impact of different personal characteristics (age, gender, cultural background or different experiences), in other cases you would like to control for these.
- You can control by eliciting these characteristics and then using these variables in regressions, or better yet control the subject group directly.
- Example: if you want to "shock" a parameter such as choice overload or frequency of donation, trying to recruit subjects with similar characteristics may make the shock have a similar impact across subjects.
- Alternatively you might actually want a varied impact in which case you might want to recruit a fixed quota of different types (such as a balance of male and female subjects).

About the Subject Choices

- For lab Experiments: subjects will be a registered pool in a university: at Warwick we have SONA, other universities might use ORSEE. https://warwick.sona-systems.com/Default.aspx?ReturnUrl=%2f

About the Subject Choices

WARWICK

- For lab Experiments: subjects will be a registered pool in a university: at Warwick we have SONA, other universities might use ORSEE. https://warwick.sona-systems.com/Default.aspx?ReturnUrl=%2f
- For online Experiments: e.g. Amazon Mechanical Turk (MTurk) or Prolific, SuveryMonkey, Qualtrics, City Surveys, etc.

About the Subject Choices

- For lab Experiments: subjects will be a registered pool in a university: at Warwick we have SONA, other universities might use ORSEE. https://warwick.sona-systems.com/Default.aspx?ReturnUrl=%2f
- For online Experiments: e.g. Amazon Mechanical Turk (MTurk) or Prolific, SuveryMonkey, Qualtrics, City Surveys, etc.
- You can also go for survey panels or attempt to gather a vocation-specific pool yourself (fund managers, politicians, etc.).

About the Subject Choices

- For lab Experiments: subjects will be a registered pool in a university: at Warwick we have SONA, other universities might use ORSEE. https://warwick.sona-systems.com/Default.aspx?ReturnUrl=%2f
- For online Experiments: e.g. Amazon Mechanical Turk (MTurk) or Prolific, SuveryMonkey, Qualtrics, City Surveys, etc.
- You can also go for survey panels or attempt to gather a vocation-specific pool yourself (fund managers, politicians, etc.).
- On top of the standard subject choice issues you also need to consider cost, feasibility and representativeness a bit more closely if you are thinking about the right pool of subjects.

In summary

WARWICK

- A good experiment identifies an interesting question or questions (issues that are better addressed through a controlled experiment than through gathering field data).
- It should determine a precise set of hypotheses.
- The design should involve a simple environment that allows you to test the hypotheses that matters.
- The more complicated the environment the more likely you are to lose control and be unable to draw inference.
- It should deal with confounding alternatives.

Good Practices

1. Pre-registration is increasingly common, though still relatively rare for lab-experiments. A common one for experimental economists is the AEA RCT registry: https://www.socialscienceregistry.org/

Good Practices

1. Pre-registration is increasingly common, though still relatively rare for lab-experiments. A common one for experimental economists is the AEA RCT registry: https://www.socialscienceregistry.org/

- The idea is to list your main objectives, measures and hypotheses in advance.
- This is good science in the sense that it prevents experimentalists from just fishing for interesting results:

  if you examine enough variables then probabilistically some will correlate.

  Or data mining afterwards to exclude certain findings.
- By pre-registering you are committing to limiting yourself to specific ideas which strengthens the results if you find anything interesting relating to your pre-registered interest.

Good Practices

WARWICK

2 Timeline

Good Practices

WARWICK

2 Timeline

- The process of starting an experimental project is long and complex. This is an idea of what you would need to think about (roughly in order):
- Formulate a research question.
- Choose design to address the research question: treatment variable(s), outcome variable(s), within vs. between, required number of independent observations, power of analysis, number of sessions/subjects, length of the study, etc.
- Prepare an experimental outline.
- Seek funding.
- Ethical approval.

Good Practices

WARWICK

2 Timeline

- The detailed instructions and design plus a pilot

Good Practices

2 Timeline

- The detailed instructions and design plus a pilot
- Write the general instruction and instructions for each treatment and an overall.
- Prepare a questionnaire (useful for controls).
- If computerized, make sure the software is capable.
- Organize money (when/how to pay).
- Recruit for the pilot.
- Run pilot experiment.

Good Practices

WARWICK

2 Timeline

- Improve the design/instructions based on the pilot experience.

Good Practices

WARWICK

2 Timeline

- Improve the design/instructions based on the pilot experience.
- At this point you might consider online pre-registration.
- Recruit subjects.
- Run the experiment.
- Analyze data and write the paper!

Good Practices

3 Good practice also involves eliminated distractions and enabling privacy.

Good Practices

3 Good practice also involves eliminated distractions and enabling privacy.

- Use the general instruction for that.
- No talking (participants, experimenters) and no distractions.
- If a subject needs help then answer questions privately and quietly.
- In smaller groups reduce the risk that one participant says something out loud and affects the whole group.
- Use appropriate screens if privacy is important (which is usually the case).

Good Practices

4 Privacy and anonymity are important if you wish to avoid social interactions and status concerns as well as for ethical reasons.

Good Practices

4 Privacy and anonymity are important if you wish to avoid social interactions and status concerns as well as for ethical reasons.

There are variations: single blind v.s. double blind (where even the experimenters cannot link data to true identity).

5 Use of deception(in Economics generally is a no-no).

Deception has big advantages (lower costs, the study of rare situations, simpler designs)

But the loss of control over beliefs and potential reputational damage are potentially very severe.

Considerations: Use of deception

Considerations: Use of deception

1. Losing the subjects trust:

Considerations: Use of deception

1. Losing the subjects trust:
   Subjects eventually realise that they have been deceived.
   In future experiments subjects will then be less likely to believe what they are told and/or
   their choices/decisions in the task is a repones to out-smarting the experimenter.

Considerations: Use of deception

1. Losing the subjects trust:
   Subjects eventually realise that they have been deceived.
   In future experiments subjects will then be less likely to believe what they are told and/or
   their choices/decisions in the task is a repones to out-smarting the experimenter.
2. This introduces problems when analyzing results.

Considerations: Use of deception

1. Losing the subjects trust:
   Subjects eventually realise that they have been deceived.
   In future experiments subjects will then be less likely to believe what they are told and/or
   their choices/decisions in the task is a repones to out-smarting the experimenter.
2. This introduces problems when analyzing results.
3. Validity of the experiment for the proposed question.

Considerations: Use of deception

1. Losing the subjects trust:
   Subjects eventually realise that they have been deceived.
   In future experiments subjects will then be less likely to believe what they are told and/or
   their choices/decisions in the task is a repones to out-smarting the experimenter.
2. This introduces problems when analyzing results.
3. Validity of the experiment for the proposed question.
4. Having a norm of no deception and being clear on this to subjects is therefore a useful norm within experimental economics.

Considerations: Use of deception

WARWICK

1. Losing the subjects trust:
   Subjects eventually realise that they have been deceived.
   In future experiments subjects will then be less likely to believe what they are told and/or
   their choices/decisions in the task is a repones to out-smarting the experimenter.
2. This introduces problems when analyzing results.
3. Validity of the experiment for the proposed question.
4. Having a norm of no deception and being clear on this to subjects is therefore a useful norm within experimental economics.
   To read more*:
   Grether, D.M. and Plott, C.R., 1979. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, 69(4), pp.623-638.

A Debate on Deception

To read more*:

- Bonetti, S., 1998. Experimental economics and deception. *Journal of Economic Psychology*, 19(3), pp.377-395.
- Hey, J.D., 1998. Experimental economics and deception: A comment. Journal of Economic Psychology, 19(3), pp.397-401.
- McDaniel, T. and Starmer, C., 1998. Experimental economics and deception: A comment. *Journal of Economic Psychology*, 19(3), pp.403-409.
- Bonetti, S., 1998. Reply to Hey and Starmer & McDaniel. *Journal of Economic Psychology*, 19(3), pp.411-414.

A Debate on Deception

WARWICK

To read more*:

- Bonetti, S., 1998. Experimental economics and deception. *Journal of Economic Psychology*, 19(3), pp.377-395.
- Hey, J.D., 1998. Experimental economics and deception: A comment. Journal of Economic Psychology, 19(3), pp.397-401.
- McDaniel, T. and Starmer, C., 1998. Experimental economics and deception: A comment. *Journal of Economic Psychology*, 19(3), pp.403-409.
- Bonetti, S., 1998. Reply to Hey and Starmer & McDaniel. *Journal of Economic Psychology*, 19(3), pp.411-414.
- Charness, G., Samek, A. and van de Ven, J., 2021. What is considered deception in experimental economics?. *Experimental Economics*, pp.1-28.
- Hertwig, R. and Ortmann, A., 2008. Deception in experiments: Revisiting the arguments in its defense. *Ethics & behavior*, 18(1), pp.59-92.

Good Practices

WARWICK

6 No pre-conceptions

Good Practices

6 No pre-conceptions

7 Be aware of drop-outs

Good Practices

6 No pre-conceptions

7 Be aware of drop-outs

8 Be aware of sample-selection bias

Good Practices

6 No pre-conceptions

7 Be aware of drop-outs

8 Be aware of sample-selection bias

9 Pre-screening and Filtering

Filter by age, subject, background, gender, nationality, language, etc.

Good Practices

6 No pre-conceptions

7 Be aware of drop-outs

8 Be aware of sample-selection bias

9 Pre-screening and Filtering

Filter by age, subject, background, gender, nationality, language, etc.

10 We can also exclude subjects who have taken part in related experiments before.

Exclude subjects with a poor show-up rate. especially important where we need a precise number of subjects.

What It Is, Why: An Overview

Four general categories of experiments:

- Laboratory experiments, e.g., Goeree & Holt (2001)*
- Artefactual Field Experiments (AFE), aka lab-in-the-field, e.g., Levitt & List (2009)*
- Framed Field Experiments (FFE) e.g., Gosnell et al. (2017)*
- Natural Field Experiments (NFE), e.g., Hallsworth et al. (2015)*

What It Is, Why: An Overview

WARWICK

	Lab	AFE
Population we study	College Students	Population of Interest
Environment	Artificial	Artificial
Type of awareness	Overt	Overt
Who do we observe?	Those Sorting into Experiment	Those Sorting into Experiment
	FFE	**NFE**
Population we study	Population of Interest	Population of Interest
Environment	Natural	Natural
Type of awareness	Overt	Covert
Who do we observe?	Those Sorting into Experiment	All in Market

How: Lab Experiment or Field Experiment

WARWICK

- Right method depends on the research question, e.g., the social situation and the treatment effects.
- Be aware of (dis)advantages and trade-offs in
  1. practical aspects and the ease of implementation
  2. inference and causal evidence
  3. "perfect" randomisation of the treatment variable (e.g., attrition, non-compliance
  4. spillover effects

- Good experiments do not (have to) mirror reality; reality is too complex

(dis)Advantages of the Laboratory

(dis)Advantages of the Laboratory

WARWICK

- The laboratory provides more control
- Typically it is easier to get strict instructions followed when experiments are run in the laboratory.
- The laboratory offers more transparency. well understood.
- In the field you may worry you use a subject pool prone to some bias, that is then attributed to the experiment.
- Laboratory experiments are more replicable.
- Ethical issues may be easier to overcome in the laboratory.
- Field experiments may simply be infeasible in terms of design, cost or opportunity.

(dis)Advantages of the Field Experiments

WARWICK

- More external validity built-in:

  The setting may seem more realistic.

  Real people vs undergraduates (but students are real people!)

  The subject pool is spot-on: use market traders to study trading strategies, use politicians to study legislative bargaining, etc.
- You may want really large samples (thousands of people)
- You may wish to test if a change would have a sizeable effect
- Sometimes the field is just the right place

How: Designing an Experiment

- Dictator game

  A simple design to elicit some measure of altruism and pro-social behaviour
- To participate Vevox from your web browser:

  `https://vevox.app/#/m/160494808`

  Session ID: 160-494-808

How: Designing an Experiment

WARWICK

- Dictator game

  A simple design to elicit some measure of altruism and pro-social behaviour
- To participate Vevox from your web browser:

  `https://vevox.app/#/m/160494808`

  Session ID: 160-494-808
- The subsequent follow-up experiments included
  - Fairness or kindness, stakes sizes, social controls, demographic, genetic variation etc. (Engel 2011)*
  - Variations of the game "Trust Game" and "Taking Game" (List 2007)*
  - An example of such studies is:

    Ariely, Bracha and Meier: "Doing good or doing well? Image motivation and monetary incentives in behaving prosocially" (AER, 2009) *

Control & Replicability

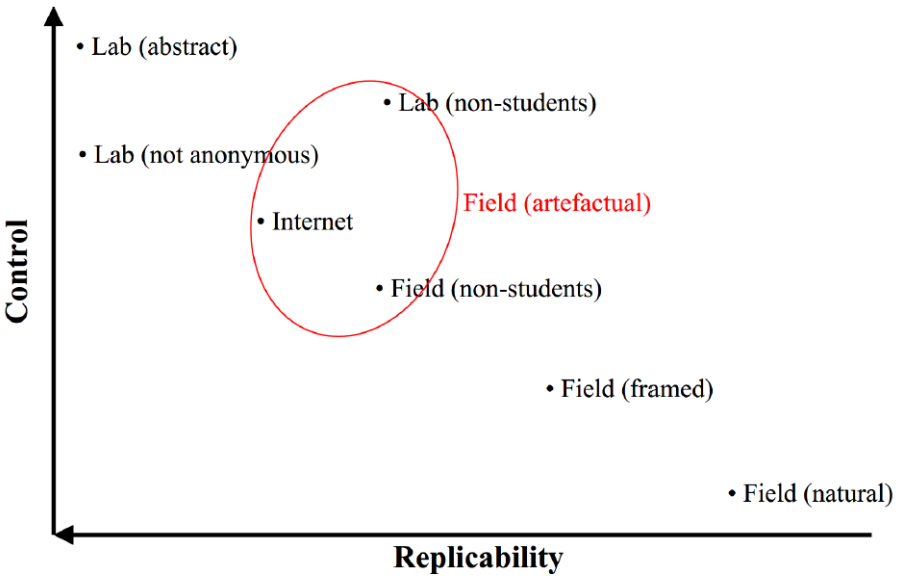


Figure: Source: from Ernesto Reuben

How: Designing An Experiment

WARWICK

- Why people give (time or money) to charities? (Vesterlund 2016)*
- Three explanations:
  1. **Pure Altruism**: they just want to help people (no strings attached!)
  2. **Impure Altruism**: they feel good about themselves after giving, aka: warm glow or joy of giving
  3. **Impure Altruism**: they do not feel good about themselves if they do not, aka, social image
- Dictator game often interpreted in terms of prosocial preferences

How: Designing An Experiment

- Ariely et al. (2009) consider the effect of visibility to test image motivation proposed by Bénabou and Tirole (2006)
- Things to consider
  1. Public v.s. Private prosocial activity
  2. Incentivised v.s. Non-incentivised prosocial activity
  3. Nature of prosocial activity, Good v.s. bad charity

How: Designing An Experiment

- Ariely et al. (2009) consider the effect of visibility to test image motivation proposed by Bénabou and Tirole (2006)
- Things to consider
  1. Public v.s. Private prosocial activity
  2. Incentivised v.s. Non-incentivised prosocial activity
  3. Nature of prosocial activity, Good v.s. bad charity
- This study used a between-subjects $2 \times 2 \times 2$ design.
- Controlled for: lab v.s field (press keys and bike for charity)
- Thought experiment: how does giving change as visibility changes, holding everything else constant?

How: Designing An Experiment, Ariely et al., 2(

WARWICK

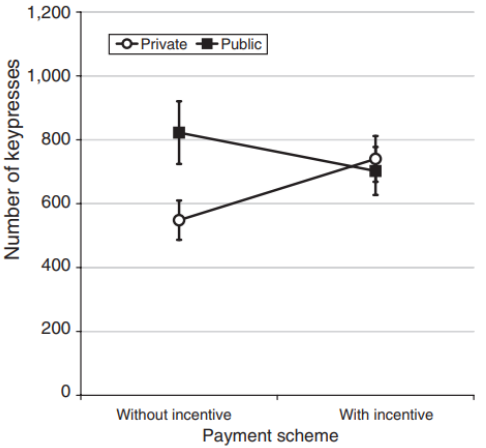
- Four treatments

How: Designing An Experiment, Ariely et al., 2(

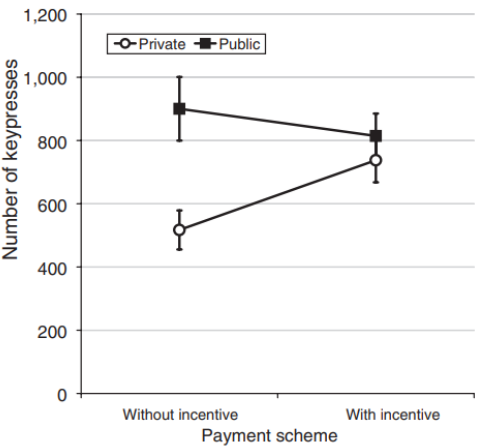
WARWICK

- Four treatments
  1. Private donation with no private incentives
  2. Public donation with no private incentives
  3. Private donation with private incentives
  4. Public donation with private incentives
- Prosocial activity is a real-effort task: the longer a subject sequentially clicks on “X” or “Z”, 1 cent goes to a charity

Findings

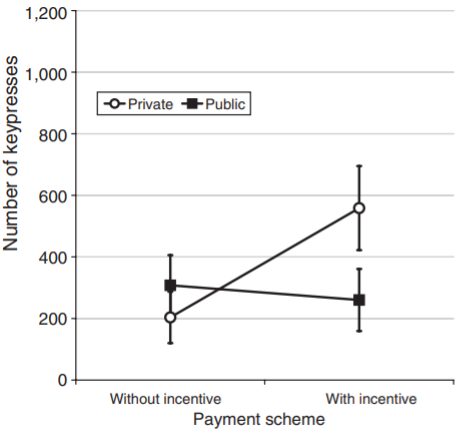


Panel A: "Good" cause

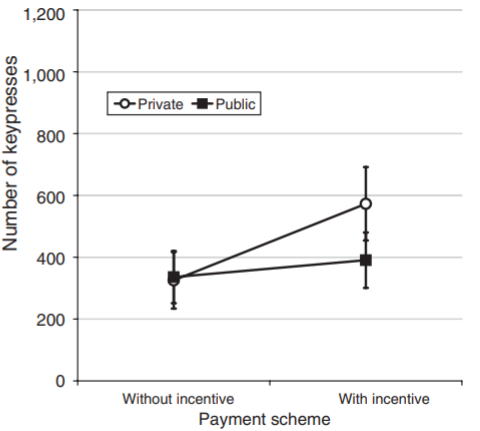


Panel B: Red Cross

Findings



Panel A: "Bad" cause



Panel B: NRA

Internal Validity

Aspects to consider:

Internal Validity

Aspects to consider:

- Correct statistical inference including internal validity

Internal Validity

Aspects to consider:

- Correct statistical inference including internal validity
  - Develop treatments that manipulate only the factor of interest, holding everything else constant
  - To identify the parameters of interest in a given design
    The Average Treatment Effect (ATE) and The Average Treatment Effect on Treated (ATT)

$$\begin{aligned} \Delta^{ATE} &= \mathbb{E}(y_i(1) - y_i(0)|\mathbf{X}) \\ \Delta^{ATT} &= \mathbb{E}(y_i(1) - y_i(0)|\mathbf{X}, \mathcal{T}_i = 1) \end{aligned}$$

  - Block everything you can and randomise otherwise
- The incentive structure of Experiments (Induced Value Theory)
- Psychologists typically do not use monetary incentives
- Emphasis on monetary incentives in experimental economics

Issues with Experiments

1. Potential lack of realism/external validity issue
2. Representativeness of subjects
3. Lack of incentives/motivation
4. Lack of clarity
5. Priming and framing

External Validity of Experimental Results

WARWICK

- External validity: Parallelism, Robustness and Inference.

  whether a causal relationship continues to hold when subjects, context, location, or treatment details are modified
- Potential threats to external validity:
  - Characteristics of the experiment
    - Experimenter-Demand Effect (Hawthorn Effect)
    - The size and source of monetary stakes
  - Selective noncompliance
  - Non-random selection into the experiment
  - Different populations

Final Note: Enhancing the Credibility of Empirical

WARWICK

- The need for replication (Camerer et al. 2016) and data Depository
- Literature reviews and meta analysis
- Eliciting Forecasts
  - Ex post, almost every experimental result is "obvious"
  - That's a huge problem for creative papers
  - One potential solution: elicit ex ante forecasts of treatment effect as natural benchmark
  - DellaVigna and Pope: "Predicting Experimental Results: Who Knows What?" (JPE, 2017)
  - DellaVigna and Pope: "What Motivates Effort? Evidence and Expert Forecasts" (REStud, 2017)

Final Note: Replication

- Sample size, and Type I and Type II errors

Truth of the alternative hypothesis	Significance of test		Total
	Significant	Not significant	
True association	$(1-\beta)\bar{y}$ [True positive]	$\beta\bar{y}$ [False negative]	$\bar{y}$
No association	$\alpha(1-\bar{y})$ [False positive]	$(1-\alpha)(1-\bar{y})$ [True negative]	$1-\bar{y}$
Total	$(1-\beta)\bar{y}+\alpha(1-\bar{y})$	$\beta\bar{y}+(1-\alpha)(1-\bar{y})$	$\bar{y}$

Note. Given the true relationship in the rows, the cells show the probability of significant (first column) and insignificant (second column) results in the literature according to the level (α) and power of the statistical tests (β) as well as the share of true associations among those investigated, $\bar{y}$.
Source: Wacholder et al. (2004, p. 440, Table 1).

Based in the table, the likelihood of a false positive among all the significant results that are reported

$$p = \frac{\alpha(1-\bar{y})}{\alpha(1-\bar{y})+(1-\beta)\bar{y}}$$

Final Note: Replication

- In judging experimental papers:
- Read the experimental instructions!
- Check how often a paper shows an effect (only once or multiple times in robustness treatments or other settings?)
- Be aware of publication and reporting biases.
- Ask yourself how many different authors have found the claimed effect

Final Note: Eliciting Forecasts

WARWICK

- This can become an important tool in trying to convince readers that your results are surprising
- Problems:
  - Only famous people can hope to get dozens of faculty members fill out their ex ante survey
  - Only really works for very simple experiments because forecasters need to read / understand experimental instructions
- If don't have access to expert forecasts, poll fellow students

Final Note

- The Voltage Effect (2022), a book by John List
- Let's say we have both internal and external validity.
- Is that enough?
- Why do so many ideas dail to deliver on their promise when scaled?

Ariely, D., Bracha, A. & Meier, S. (2009), 'Doing good or doing well? image motivation and monetary incentives in behaving prosocially', American economic review **99**(1), 544–55.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T. et al. (2016), 'Evaluating replicability of laboratory experiments in economics', Science **351**(6280), 1433–1436.

Engel, C. (2011), 'Dictator games: A meta study', Experimental economics **14**(4), 583–610.

Goeree, J. K. & Holt, C. A. (2001), 'Ten little treasures of game theory and ten intuitive contradictions', American Economic Review **91**(5), 1402–1422.

Levitt, S. D. & List, J. A. (2009), 'Field experiments in economics: The past, the present, and the future', European Economic Review **53**(1), 1–18.

List, J. A. (2007), 'On the interpretation of giving in dictator games', Journal of Political economy **115**(3), 482–493.

Nikiforakis, N. & Slonim, R. (2019), 'Editors' preface: Trends in experimental economics (1975–2018)'.

Reuben, E., Li, S. X., Suetens, S., Svorenčík, A., Turocy, T. & Kotsidis, V. (2021), 'Trends in the publication of experimental economics articles', Available at SSRN 3907528 .

Vesterlund, L. (2016), 2. Using Experimental Methods to Understand Why and How We Give to Charity, Princeton University Press, pp. 91–152.
URL: *https://doi.org/10.1515/9781400883172-003*