Quantitative & Qualitative Methods for Business

Notes for guided private self-study hours

© Prof. Dr. Stephan Huber (Stephan.Huber@hs-fresenius.de)

Last compiled on 22 February, 2023

Contents

Preface

Copyright

Message on ILIAS

Dear students,

in this studymag you find a catalog of tasks that relate to the content of the module "Quantitative & Qualitative Methods for Business".

Please contact the subject instructor to clarify which sections he/she recommends for guided self-study.

Solution outlines are provided for all exercises.

A pdf version of the studymag is available here: PDF Version

If you find that you have gaps in your understanding at one point or another while working through these reflection questions, you are welcome to bring any questions you may have to the lecture.

Figure 1: Prof. Dr. Stephan Huber

About the author[1]

Prof. Dr. Stephan Huber is Professor of International Economics and Data Science at HS Fresenius and holds a Diploma in Economics from the University of Regensburg and a Doctoral Degree (summa cum laude) from the University of Trier. He completed postgraduate studies at the Interdisciplinary Graduate Center at the Institute for Labor Law and Industrial Relations in the European Union (IAAEU) in Trier. He was a research assistant to Prof. Dr. Dr. h.c. Joachim Möller at the University of Regensburg, post-doc at the Leibniz Institute for East and Southeast European Studies (IOS) in Regensburg and freelancer at Charles University in Prague.

He has worked as a lecturer at various institutions including the TU Munich, the University of Regensburg, Saarland University, and the Universities of Applied Sciences in Frankfurt and Augsburg. He has also taught abroad for the University of Cordoba in Spain and the University of Perugia. Professor Huber has published his work in international journals such as the Canadian Journal of Economics and the Stata Journal. More on his work can be found on his private homepage www.t1p.de/stephanhuber.

Introduction

Welcome to the self-study module of the course "Quantitative & Qualitative Methods for Business". In the course, you will consolidate and enhance your research methodological skills. This Studymag is designed to help you develop your ability to make evidence-based decisions, regardless of the instructor's course history.

In today's data-driven business environment, it is more important than ever

[1]Picture is taken from https://sites.google.com/view/stephanhuber

to be able to situate practical problems in scientific contexts, generate reliable arguments from scientific discourse, critically evaluate research findings, and interpret and present data.

By the end of this module, you will be able

- to assess and discuss a research designs and strategies, respectively,
- to explain a selective set of quantitative and qualitative methods for collecting, analyzing, and interpreting data, and distinguish and
- to discuss some empirical strategies for identifying causal mechanisms, causes, and effects.

We are excited to support you as you acquire these valuable skills and we look forward to seeing your progress throughout the course.

Personal note

Dear students,

If the title of this course "Quantitative & Qualitative Methods for Business" seems uninteresting to you, I assure you that it is actually quite exciting because it focuses on how we can use information to understand how the world and business works and how to interpret facts. The course will enhance your data literacy, help you think critically, and improve your personal decision-making skills.

One way we can do this is by understanding the differences between quantitative and qualitative data and how they can be used to inform our choices.

Quantitative data is information that can be measured, such as numbers and statistics, while qualitative data is information that cannot be measured and is often expressed in words or other non-numerical forms.

Both forms of information are crucial for making good decisions. Without sufficient information, it can be difficult to evaluate the options and potential outcomes of a decision, leading to poor or uninformed choices. In general, the more information a decision-maker has and the faster and better the information can be used, the better they will be to make a sound decision.

The methods we discuss in this course will help you systematically gather information and make sense of it.

Enjoy the course and especially this self-study guide that will help you learn for about 21 hours.

Chapter 1

Doing research

Figure 1.1: Causal Inference: The Mixtape. A textbook by Scott Cunningham (2021).

> *"True scientists do not collect evidence in order to prove what they want to be true or what others want to believe. That is a form of deception and manipulation called propaganda, and propaganda is not science. Rather, scientific methodologies are devices for forming a particular kind of belief. Scientific methodologies allow us to accept unexpected, and sometimes undesirable, answers."* (Cunningham 2021, 10)

1.1 Everybody can do research

Research often involves exploring unknown territory and seeking out new information through methods such as attending conferences, conducting interviews and experiments, and reading related research. This process can lead to the discovery of valuable techniques or insights that address important issues in society or science.

Figure 1.2: Zora Neale Hurston, 1891-1960[1]

> *"Research is formalized curiosity. It is poking and prying with a purpose."* (Hurston 2010)

Before I go into how empirical research can and should be conducted, I would like to assert that each of us is a researcher in some sense and that you don't need a degree or a higher education to be a (good) researcher. Each of my four children (ages 2, 5, 6, and 8), for example, explores the world and learns something new every day. Even though none of my children is yet able to verify the novelty of their acquired knowledge and write it down in scientific form, I will claim that mine, like practically all children, are already little scientists. Why? Well, they explore unknown territory and search for information to discover new techniques that will make their lives pleasant. Of course, they don't attend conferences or read journals to do this. They have never heard terms such as ontology, epistemology, axiology, or quantitative and qualitative methods. They are using methods that they have mastered for their age. They interview me, my wife and all other people around and they conduct experiments. For example, all my children liked to throw plates,

[1]Photography is taken from Library of Congress: Prints & Photographs Division, Carl van Vechten Collection, Reproduction Number LC-USZ62-54231, see: https://www.loc.gov/pictures/item/2004663047/..

cutlery, cups and alike from the table when they were about one year old. At first the throwing was just an accident, but they quickly found out that each throw was followed by a sound when the object touched the stone floor. My first son, in particular, took great delight in making these sounds. He threw everything within reach to the ground and giggled with joy at the clink he made when the object hit the ground. Perhaps he was also enjoying the attention he was getting from us parents through these actions. In any case, the behavior annoyed us. Wiping food scraps off the floor is not a nice thing to do. Unfortunately, at that time my son did not accept any argument to refrain from throwing. Neither a stern look nor a definite "no" helped to stop this behavior. Too great was the joy at the relationship he had figured out, which was, "I throw something off the table and it always clangs beautifully loud." So I started to do some research to figure out what I could do to stop him. The short answer I found can be summed up pretty well as "nothing". There is practically no good method to change the behavior without possibly negatively influencing his early childhood development. The reason is he did some research and we should not suppress that. Besides nature and material research he did social research: He found out that things fall to the ground (gravity), that things break and make different sounds (material research), and that other people notice him when he throws things (social research).

Figure 1.3: Children as little researcher[2]

[2]Image by macrovector on Freepik, see: https://www.freepik.com/free-vector/

Once, when we were eating at a friend's house, my son (once again) threw everything off the table one after the other in unobserved moments. This time, however, it made no noise. The carpet under the table muffled everything. My son was irritated and at some point became really angry. Why? Well, his surely believed reality and his law "I throw something from the table and then it always clangs beautifully loud." was falsified. Soon he understood that his law only had to be adapted a little. It was then: "I throw something from the table and it clangs then beautifully loudly if a stone floor is under me." He repeated his experiments for a few more weeks, to check its validity. In the meantime he does other experiments trying to contribute to his own knowledge.

In general, the purpose of research is to find new knowledge or discover new ways to use existing knowledge in a creative way so as to generate new concepts, methodologies, inventions and understandings that -now or later- may be of some value for the human mankind. In simple terms, we aim to find something out. We aim to find a new law, a new relationship, a new insight. Or, we aim to challenge and revise existing insights on how the world works. You don't need a degree to do that. All you need is interest, open-mindedness, and a willingness to revise your ideas about how the world works. The latter is perhaps the most important skill you need to be a good researcher. Otherwise, one is a narrow-minded, and bigoted person who is too proud to follow up an insight with a change of mind.

I myself have a quick and happy tendency to change my views because it is a statement of a fresh understanding. Here are two more quotes from historically more significant people than me that are along the same lines and should convince you that changing your mind is not a sign of weakness, but of strength. Especially in science, the willingness to change one's mind is essential.

> *"When the facts change, I change my mind. What do you do, sir?"*[4]

`kindergarten-set-isolated-icons-with-toys-characters-kids-practicing-with-teacher-playing-games-vector-illustration_26760074.htm`

[3]Photography is public domain and stems from https://de.wikipedia.org/wiki/John_Maynard_Keynes#/media/Datei:Keynes_1933.jpg

[4]This quote is often attributed to Keynes, but there is no clear evidence for it, see: https://quoteinvestigator.com/2011/07/22/keynes-change-mind/

Figure 1.4: John Maynard Keynes (1883-1946)[3]

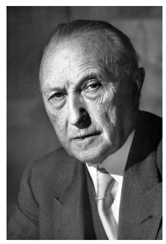

Figure 1.5: Konrad Adenauer (1876-1967)[5]

> *"What do I care about the rubbish I said yesterday? No one can stop me from getting smarter every day."* ("Was interessiert mich mein Geschwätz von gestern? ... es kann mich doch niemand daran hindern, jeden Tag klüger zu werden.")[6]

1.2 It's difficult to do good research

Simply trying something and seeing what happens, like my children do, is a research method that relies on luck and chance. Before I go into more grown-up ways of doing research, I want to emphasize that the role of chance and serendipity in research is often downplayed and not acknowledged. The most well-known example of such research is the discovery of penicillin by Alexander Fleming. In 1928, Fleming was studying the properties of staphylococcus bacteria when he noticed that a mold called Penicillium notatum had contaminated one of his bacterial cultures. He noticed that the mold seemed to be inhibiting the growth of the bacteria, and he began to

[5]This photography from 1952 is public domain and stems from the Bundesarchiv, B 145 Bild-F078072-0004, Katherine Young, CC BY-SA 3.0 DE.

[6]Freely quoted (and translated) from Weymar (1955, 521)

investigate this further. Eventually, he was able to isolate and purify the active ingredient in the mold, which he named penicillin, and he discovered that it had powerful antibiotic properties. This discovery revolutionized the field of medicine and has saved countless lives.

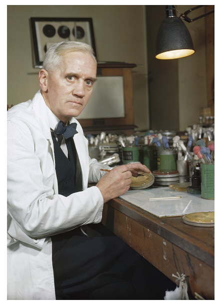

Figure 1.6: Sir Alexander Fleming (1881-1955)[7]

Doing something on purpose and observing how things respond to the action can be considered a research strategy. Acting like a child or just waiting for something to happen by chance can also be considered a research strategy, and of course this can contribute greatly to knowledge. However, it are a naïve and poorly targeted strategies to conduct research. There are more grown-up research methods that are targeting more precisely the gaps in our knowledge and speed up innovation in the field where progress is desperately needed.

What we do and how we observe what is happening should be done in a way to increase the likeliness that we find something new and interesting and in a way that allows us to be rather certain that our results are true and are less likely to be falsified soon later.

For example, assume that there is a disease that can kill people. The childish way to figure out how to cure the disease would be to observe who gets sick and who dies, and finally hope to find a cure for the disease by accidental observation. This is most likely not a very promising and quick method. It would probably be much better to study the matter systematically. For example, a laboratory should first seek to isolate the causative virus or bacterium in order to be able to grow and study it outside the danger to

[7] Photography is public domain and stems from https://en.wikipedia.org/wiki/File:Synthetic_Production_of_Penicillin_TR1468.jpg

humans. Once this is done, we need a precise plan on how we can use all the available knowledge to cure the disease, protect people from infection, or help them survive the disease. In short, we need a strategic way to conduct research, i.e., a research strategy or design.

A *research strategy* is a general plan for conducting a study and a *research design* is a detailed plan for conducting the study. These words are frequently used interchangeably. A research strategy depends on many things including the question, the resources available, the current state of knowledge, the ambitions, whether quantitative or quantitative data are used, and what is considered to be the criteria of good research.

Before discussing some research strategies that can provide reasonable answers to certain types of questions, we should clarify how to ask a research question and what qualifies a research question.

1.3 Asking questions lika a good researcher

Unfortunately, there is no one research strategy that is appropriate for all questions and, what is worse, there is still controversy about what constitutes good research and how to properly ask a research question. In particular, this controversy takes place between researchers who use quantitative data and statistical methods and researchers who use qualitative data and methods.

Quantitative researchers are more interested in determining the causes of effects using experiments with measurable results, and they attempt to quantify the effects of causes. Qualitative researchers also try to determine the causes of effects . However, their data analysis does rely less on statistical inference. A qualitative data set not necessarily requires (large) random samples or structured data (all the data that you can structure in a spreadsheet) in general, but allows to analyze selective and unstructured data (that is data in form of audio, video, text, images and alike). Qualitative research methods allow to classify these data into patterns or to interpret them in a meaningful way in order to arrive at results. Qualitative researchers are more concerned with the *why and how* of decision making and examine people's behavior, beliefs, perceptions of events, experiences, attitudes, interactions, and more in great depth.

In empirical research, inductive and deductive are two different approaches to reasoning. *Inductive reasoning* is a process of collecting data from various sources, such as interviews, surveys or observations, and then use this data to identify patterns, themes, or relationships that can form the basis of a new hypothesis or theory. The goal of these exploratory studies, is to generate new ideas or insights about a topic, rather than testing a specific hypothesis. *Deductive reasoning* is a process in which the researcher starts with a general theory or hypothesis with the goal to test a specific hypothesis or theory. In most cases, a combination of both inductive and deductive reasoning may be used to formulate the research question and to design the empirical identification strategy.

In what follows, however, we focus on the criteria for *good research* that are more commonly used in evaluating the quality of quantitative research.

Exercises (The Effect ch.1+2)

Read chapter 1 and 2 of Huntington-Klein (2022) and answer the questions below. The book is freely available at https://theeffectbook.net and here is the link to chapter 1: https://theeffectbook.net/ch-TheDesignofResearch.html

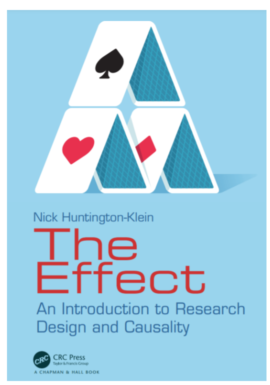

Figure 1.7: The book *The Effect: An Introduction to Research Design and Causality* by Nick Huntington-Klein (2022).

1. What is the main focus of the book the author is writing about?
    a) Philosophy of science
    b) Qualitative research methods

   c) Empirical research and quantitative methods to identify and measure causal effects
   d) Statistics
2. What is the main challenge faced by quantitative empirical research, according to the author?
   a) Difficulty in obtaining accurate measurements
   b) Difficulty in interpreting measurements
   c) Difficulty in obtaining data that allows to answers the research question
   d) Difficulty in designing a research that gets a lot of attention
3. What is the author's main point about research questions?
   a) They should be well-defined, answerable, and understandable
   b) They should be simple and easy to answer
   c) They should be related to the world of traffic
   d) They should be related to the field of quantum mechanics

Answers: 1. c), 2. c), 3. a)

1.4 Features of good research

In order to make you a competent researcher who does not have to wait for a lucky chance but has a clear strategy, let's discuss the criteria of a good research. Before I do that, however, I must make a disclaimer: there is a lack of consensus on what constitutes high-quality research in social sciences. In my experience, the practical benefits of such a tedious discussion are quite small. All I like to put forward is that I believe that all social science disciplines such as sociology, anthropology, psychology, economics, business administration, and education using quantitative methods agree that good research should be replicable, reproducible, transparent, reliable, and valid.

1.4.1 Reliability and validity

A research design is a plan to examine information in a systematic and controlled way so that the results of the research are *valid and reliable*.

Validity refers to the accuracy and truthfulness of research findings. In other words, if a study is valid, it should measure what it is intended to measure and produce results that are representative of the population being studied.

Validity is important because it helps to ensure that the conclusions drawn from a study are supported by the data and are not based on flawed or biased methods.

Reliability refers to the consistency and stability of research findings. In other words, if a study is reliable, it should produce similar results if it is repeated using the same methods and conditions. Reliability is important because it helps to ensure that the results of a study are not simply due to chance or random error.

Both reliability and validity are important considerations in research, and researchers strive to maximize both in their studies. However, it is important to note that it is often difficult to achieve both at the same time, and trade-offs may need to be made between the two.

Statement: A good research design should aim to minimize bias and maximize the reliability and validity of the research. It should also be appropriate for the research question being asked and the resources available to the researcher.

High reliability and low validity

An example of a study that has high reliability but low validity is a study that measures the weight of a group of people using a digital scale. If the scale is consistently accurate and produces the same weight measurements each time it is used, then the study has high reliability. However, if the scale is not calibrated correctly and produces inaccurate weight measurements, then the study has low validity.

Another example of a research design that has high reliability but low validity is a study that uses a highly reliable measurement tool, such as a standardized test, to measure a concept that is not directly related to the research question being asked. For example, a study that uses a standardized math test to measure students' critical thinking skills may have high reliability because the test is consistently accurate and produces similar scores each time it is administered. However, the study may have low validity because the math test is not an appropriate tool for measuring critical thinking skills. As a result, the results of the study may not be representative of the students' true critical thinking abilities.

High validity and low reliability

An example of a study that has high validity but low reliability is a study that asks people to self-report their eating habits. While the study may produce accurate and representative results about people's eating habits, the self-reported data may vary from person to person and may not be consistent over time. As a result, the study has high validity but low reliability.

Another example of a study that has high validity but low reliability is a study that uses a highly valid measurement tool, such as a survey, to measure a concept that is directly related to the research question being asked. However, the study may have low reliability because the survey is not administered consistently or the responses are not accurately recorded. For example, a study that uses a survey to measure students' attitudes towards school may have high validity because the survey is relevant to the research question and accurately measures the students' attitudes. However, if the survey is not administered consistently or the responses are not accurately recorded, the study may have low reliability. As a result, the results of the study may not be representative of the students' true attitudes towards school.

Trade-offs between reliability and validity

In research design, trade-offs may need to be made between reliability and validity. For example, a study that uses a highly reliable measurement tool may not be valid if the tool is not appropriate for the research question being asked. Similarly, a study that uses a highly valid measurement tool may not be reliable if the tool is prone to producing inconsistent results. As a result, researchers must carefully consider both reliability and validity when designing a study and make trade-offs as necessary to maximize the overall quality of the research.

1.4.2 Generalizability

Coming back to my little son who threw everything within reach to the ground and giggled with joy at the clink he made when the object hit the ground. He identified a cause-and-effect relationship through an experiment in an controlled environment. His law "I throw something off the table and it always clangs" worked in our home. To our regret, it was replicateable and he really tried hard to falsify it. Moreover, his study was reasonable valid as his study design, conduct, and analysis could answer his research questions without bias (at least ignoring the other noises that his sibling and parents make coincidentally during his experiment). Scientist call this *internal validity*. However, he also found out that when he leaves our home, things are sometimes a bit different, for example, if there is a carpet under the table. Thus, his insights from our home findings can't be generalized to other contexts, at least not without further specifications. Scientist call this *external validity*.

Statement: Internal validity examines whether the study design, conduct, and analysis answer the research questions without bias. External validity examines whether the study findings can be generalized to other contexts.

1.4.3 Replicability, reproducibility, transparency, and other criteria

It must be possible to repeat the research conducted for several reasons. For example, if you can repeat a study with slightly changed parameters, you are able to improve its external validity and show that the conclusions drawn are reliable. To be able to repeat a study, everything that is important for drawing a conclusion from the research has to be mentioned. This is what we call *transparency*. Moreover, everything in the study must have been done in such a way that we can check the results for truth. In the best case, it is possible to *reproduce* the results in the same way they were obtained in the study. Sometimes this is not possible because, for example, we can never really ask the same people again in a survey, and even if we found the same people, they would have gotten older and not be the same people as before. In such a case, it should at least be possible to *replicate* the research. This means that we can basically do the same thing in a setting that differs only in those things that we cannot avoid to be different. For example, by interviewing a group of people who match the people in the study to replicate them on all the important characteristics like age.

In an empirical quantitative research study, for example, the data and the code written to process the data and analyze it should be accessible to everyone.

In a qualitative study, all sources of information should be stated, and the circumstances leading to a conclusion should be fully explained. For example, all transcripts of interviews conducted should be made available.

The researcher should provide rich and detailed descriptions of the data and the context in which it was collected. Research should be provided with rich, nuanced, and multi-layered accounts of social phenomena by describing and interpreting the meanings, beliefs, and practices of the people being studied. That is known as *thick description.* Researchers typically employ a variety of methods such as participant observation, in-depth interviews, and document analysis, and they often use multiple sources of data to triangulate their findings. The goal is to provide a holistic and broad understanding of the phenomenon being studied, rather than a narrow view from the researcher's perspective.

There are some other criteria of good research that are worth mentioning:

Credibility

The research should be trustworthy and believable, and the researcher should provide detailed descriptions of the methods used to ensure transparency.

Reflective Practice

The researcher should engage in reflexive practice throughout the research process, which means to be critically aware of oneself, one's own assumptions, and one's own role in the research process.

Triangulation

The researcher should use multiple methods, sources, and perspectives to increase the credibility of the findings (also see *thick description* above).

Transferability

The conclusions drawn from looking at mostly unstructured data in qualitative research can hardly be generalized in a strict sense, as they depend crucially on the context of the object of study. For example, generalizability is essentially impossible in a qualitative case study, since everything depends on the specific situation of an individual, a company, or a group of people considered in the specific setting. This means that in a case study or interview, we may be looking at only a few or even a single observation that cannot be considered *representative* of the larger population, as generalizability does. Transferability, on the other hand, gives the reader the ability to transfer the findings into other contexts. The ability to transfer contextual findings to other cases is a goal of qualitative research, and the author of a study should attempt to offer the information in a way that allows the reader to transfer the findings to the setting or situation with which he or she is familiar.

1.5 The role of resources, data and ethics

There are several types of research designs, including experimental designs, quasi-experimental designs, and observational designs. Each of these designs took advantage of various empirical methods and statistical procedures. We will discuss some of them later on. The choice of research design, of course, should depend on the research question being asked, the resources available, and the type of data that is being collected. The research design should also take into account any ethical considerations that may be relevant to the research. The research design should be chosen so that it is well suited to answer the research question. For example, if one is interested in the question "Why do some people get sick with a certain disease and others do not?" then an observational study design to determine possible *causes of effect* may be appropriate. These identified potential causes should then be verified followed by an experimental study. Relatively, a statistical analysis should be used which would allow the *effects of causes* to be evaluated. The aim should be to identify necessary and sufficient circumstances to develop a disease. Also circumstances should be described that favor a disease.

If the question is a "how" question, for example, "How do parents feel when their child throws everything off the table?" then interviews might be an appropriate study design. If available resources such as time, funding, and staff are limited, you might also consider conducting an (online) survey in which parents are asked standardized questions about their feelings. In any way, the chosen research design must be feasible given the resources available.

In answering a question, a researcher should know, state, and discuss all the assumptions and unexamined beliefs that led him to his conclusion. However, since resources for conducting and explaining research are limited, special attention should be paid to what are called *critical assumptions*. These are assumptions that must be true in reality, otherwise the research is meaningless. Therefore, researchers should make great efforts to identify and validate these assumptions.

The type of data that is being collected is another important factor to consider when choosing a research design. Different types of data, such as quantitative data, qualitative data, or a combination of both, may require different methods of collection and analysis. For example, quantitative data, such as numerical data, can be collected through methods such as surveys and analyzed using statistical techniques, whereas qualitative data, such as interview transcripts, may require more interpretive methods of analysis.

Finally, the researcher should also take into account any ethical considerations that may be relevant to the research. For example, if the study involves human subjects, the researcher must ensure that the study is conducted in accordance with ethical principles such as informed consent and confidentiality. Additionally, the researcher should ensure that the potential benefits of the study outweigh any potential risks to the subjects.

Exercises (features of research)

1. Which of the following best defines reliability in research?
    a) The extent to which a measurement tool produces consistent results
    b) The extent to which a study's results accurately reflect the concept being measured
    c) The extent to which a study's results can be generalized to other populations
    d) The extent to which a study's results are statistically significant
2. Which of the following best defines validity in research?
    a) The extent to which a measurement tool produces consistent results
    b) The extent to which a study's results accurately reflect the concept being measured
    c) The extent to which a study's results can be generalized to other populations
    d) The extent to which a study's results are statistically significant
3. Which of the following is an example of a study with high reliability but low validity?
    a) A study that uses a highly reliable measurement tool to measure a concept that is directly related to the research question being asked
    b) A study that uses a highly valid measurement tool to measure a

concept that is not directly related to the research question being asked
   c) A study that uses a highly reliable measurement tool to measure a concept that is not directly related to the research question being asked
   d) A study that uses a highly valid measurement tool to measure a concept that is directly related to the research question being asked

4. Which of the following is an example of a study with high validity but low reliability?
   a) A study that uses a highly reliable measurement tool to measure a concept that is directly related to the research question being asked
   b) A study that uses a highly valid measurement tool to measure a concept that is not directly related to the research question being asked
   c) A study that uses a highly reliable measurement tool to measure a concept that is not directly related to the research question being asked
   d) A study that uses a highly valid measurement tool to measure a concept that is directly related to the research question being asked
5. What does internal validity examine in a study?
   a) The ability to replicate the study
   b) The generalizability of the study's findings
   c) Whether the study design, conduct, and analysis answer the research questions without bias
   d) All of the above
6. What does external validity examine in a study?
   a) The ability to replicate the study
   b) The generalizability of the study's findings
   c) Whether the study design, conduct, and analysis answer the research questions without bias
   d) None of the above
7. What is transparency in research?
   a) The ability to replicate a study
   b) The generalizability of the study's findings
   c) The availability and accessibility of the data and materials used in

a study for others to review
    d) The ethical considerations of the research
8. What are the different types of research design discussed in the text?
    a) Experimental designs, quasi-experimental designs, and observational designs
    b) Experimental designs and descriptive designs
    c) Quasi-experimental designs and observational designs
    d) None of the above
9. Why is replicability important in a study?
    a) To be able to repeat a study with slightly changed parameters and thus improve the external validity
    b) To be able to check the results of the study for truth.
    c) To be able to reproduce the results in the same way they were obtained in the study
    d) All of the above

Answers:

1. a), 2. b), 3. c), 4. d), 5. c), 6. b), 7. c), 8. a), 9. d)

Glossary

- **Generalizability:** The extent to which the results of a study can be applied to other populations or contexts.
- **Internal validity:** The degree to which a study's results can be attributed to the specific variables or factors being studied, and not to other extraneous factors.
- **External validity:** The degree to which a study's results can be generalized to other populations or contexts outside of the specific sample or setting of the study.
- **Quantitative data:** Data that can be measured and quantified.
- **Qualitative data:** Data that cannot be easily measured or quantified.
- **Quantitative research:** A research approach that uses statistical methods and experiments to determine the causes of effects, to quantify the effects of causes, or to describe data.
- **Qualitative research:** A research approach that uses unstructured data and methods to examine, for example, people's behavior, beliefs, and experiences in depth, rather than quantifying results.

- **Reflective Practice:** A form of self-evaluation used to analyze one's own thoughts and actions.
- **Reliability:** The consistency of a study's results to produce similar results when repeated.
- **Research design:** A detailed plan for conducting a study, frequently used interchangeably with research strategy.
- **Research method:** A procedure used to conduct a study or investigation to gain knowledge or understanding about a particular topic.
- **Research question:** A question or problem that a study aims to answer or solve.
- **Research strategy:** A general plan for conducting a study, frequently used interchangeably with research design.
- **Replicability:** The ability of a study to be repeated with new data.
- **Reproducibility:** The ability of a study to be repeated and produce the same results, often used interchangeably with replicability.
- **Serendipity:** The role of luck and unexpected events in research.
- **Thick Description:** A detailed narrative used to explain a situation and its context.
- **Credibility:** A quality criterion in qualitative research, which refers to confidence in the truth value of the data and interpretations of them.
- **Transparency:** The degree to which a study's methods and data are easily accessible and understandable to others, allowing for the study to be independently evaluated and replicated.
- **Triangulation:** A method used in qualitative research to verify the accuracy of data by combining multiple sources of information.
- **Validity:** The degree to which a study measures what it is intended to measure, and the extent to which the results of the study can be considered accurate and meaningful.
- **Structured data:** Data that can be easily organized and analyzed in a structured format, such as a spreadsheet.
- **Unstructured data:** Data that cannot be easily organized and analyzed in a structured format, such as text, images, and audio.

Chapter 2

Identification

In empirical research, *identification refers to the process of establishing a clear and logical relationship between a cause and an effect.* This involves demonstrating that the cause is responsible for the observed effect, and that there are no other factors that could potentially explain the effect. The goal of identification is to provide strong evidence that a particular factor is indeed the cause of a particular outcome, rather than simply coincidentally happen. In order to identify a cause-and-effect relationship, researchers can use experimental or non-experimental, i.e., observational data, or both. The next section discusses how to generate or collect this data.

2.1 How to get data

There are several ways to get data which allows you to identify a cause-and-effect relationship:

2.1.1 Interviews

An interview is normally a one-on-one verbal conversation. Interviews are conducted to learn about the participants' experiences, perceptions, opinions, or motivations. The relationship between the interviewer and interviewee must be taken into account and other circumstances (place, time, face to face,

[1]The picture is free to use under the Pixabay license, see: https://pixabay.com/images/id-4799256/

Figure 2.1: Interview[1]

email, etc.) should be taken into account. There are three types of interviews structured, semi-structured, and unstructured. *Structured interviews* use a set list of questions and hence are like a verbal surveys. In *unstructured interviews* the interviewer doesn't use predetermined questions but only a list of topics to address. Semi-structured interviews are the middle ground. Semi-structured interviews require the interviewer to have a list of questions and topics pre-prepared, which can be asked in different ways with different interviewee/s. Semi-structured interviews increase the flexibility and the responsiveness of the interview while keeping the interview on track, increasing the reliability and credibility of the data. Semi-structured interviews are one of the most common interview techniques.

Structured interviews use a predetermined list of questions that must be asked in a specific order, improving the validity and trustworthiness of the data but lowering respondent response. Structured interviews resemble verbal questionnaires. In unstructured interviews, the interviewer has a planned list of subjects to cover but no predetermined interview questions. In exchange for less reliable data, this makes the interview more adaptable. Long-term field observation studies may employ unstructured interviews. The middle ground are interviews that are semi-structured. In semi-structured interviews, the interviewer must prepare a list of questions and themes that can be brought up in various ways with various interviewees.

Interviews allow you to address a cause-and-effect relationship fairly directly, and it can be a good idea to interview experts and ask some *why* and *how* questions to gather initial knowledge about a particular topic before further elaborating your research strategy. For example, I interviewed kindergarten teachers with many years of experience working with children, as well as other parents, to get information on how to solve the problem of my children throwing plates around the dining room. However, findings based on interviews are not very valid or reliable because the personal perceptions of both the interviewer and the interviewee can have an impact on the conclusions drawn. For example, I received very different tips and explanations because of the personal experiences of the people I interviewed. Unfortunately, I could not really ask my son why he was misbehaving. His vocabulary was too limited at the time, and even if he could speak, he would probably refuse to tell me the truth.

2.1.2 Surveys

Figure 2.2: Survey[2]

In contrast to an interview a survey can be sent out to many different people. Surveys can be used to identify a cause-and-effect relationship by asking questions about both the cause and the effect and examining the responses. For example, if a researcher wanted to determine whether there is a relationship between a person's level of education and their income, they could conduct a survey asking participants about their education level and their income. If the data shows that participants with higher levels of education tend to have higher incomes, it suggests that education may be a cause of higher income. However, it is important to note that surveys can only establish a correlation between variables, but it is difficult to claim that correlations that where found through the survey imply a causal relationship. To establish a causal relationship, a researcher would need to use other methods, such as an experiment, to control for other potential factors that might influence the relationship that the respondent does not see.

2.1.3 Case studies

[2]The picture is free to use under the Pixabay license, see: https://pixabay.com/images/id-1594962/

[3]Picture is free to use and stems from RODNAE Productions and was taken from Pexels: https://www.pexels.com/de-de/foto/linse-geschaft-papier-aufsicht-7947753/

Figure 2.3: Case study[3]

Case studies involve in-depth examination of a single case or a small number of cases in order to understand a particular phenomenon. Case studies can be conducted using both quantitative and qualitative methods, depending on the research question and the data being analyzed. While it is reasonable to find causal effects in the particular case, it is problematic to generalize the causal relationship. To establish a general causal relationship, a researcher would need to use other methods, such as an experiment, to control for other potential factors that might influence the relationship that the respondent does not see.

2.1.4 Experiments

One way to clearly identify a cause-and-effect relationship is through experiments, which involve manipulating the cause (the independent variable) and measuring the effect (the dependent variable) under controlled conditions (we will later on define precisely what is meant here). Experiments can be conducted using both quantitative and qualitative methods. Here are some examples:

- A medical study in which a new drug is tested on a group of patients, while a control group receives a placebo.
- An educational study in which a group of students is taught a new method of learning, while a control group is taught using the traditional method.
- An agricultural study in which a group of crops is treated with a new fertilization method, while a control group is not treated.
- A study to determine the effect of a new training program on employee productivity might involve randomly assigning employees to either a control group that does not receive the training, or an experimental group that does receive the training. By comparing the productivity

of the two groups, the researchers can determine if the new training program had a causal effect on employee productivity.

- A study to determine the effect of a new advertising campaign on sales might involve randomly assigning different groups of customers to be exposed to different versions of the campaign. By comparing the sales of the different groups, the researchers can determine if the advertising campaign had a causal effect on sales.
- In *experimental economics*, experimental methods are used to study economic questions. In a lab-like environment data are collected to investigate the size of certain effects, to test the validity of economic theories, to illuminate market mechanisms, or to examine the decision making of people. Economic experiments usually motivates and rewards subjects with money. The overall goal is to mimic real-world incentives and investigate things that cannot be captured or identified in the field.
- In *behavioral economics*, laboratory experiments are also used to study decisions of individuals or institutions and to test economic theory. However, it is done with a focus on cognitive, psychological, emotional, cultural, and social factors.

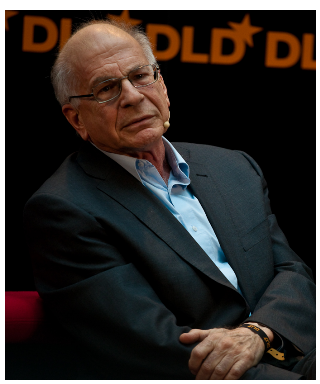

Figure 2.4: Daniel Kahneman (*1934)[4]

In 2002 the Nobel Prize of Economics was awarded to Vernon L. Smith "for having established *laboratory experiments* as a tool in empirical economic analysis, especially in the study of alternative market mechanisms" (The Royal Swedish Academy of Sciences 2002) and Daniel Kahneman "for having

[4]The picture of Kahneman is free to use and stems from https://commons.wikimedia.org/wiki/File:Daniel_Kahneman_(3283955327)_(cropped).jpg

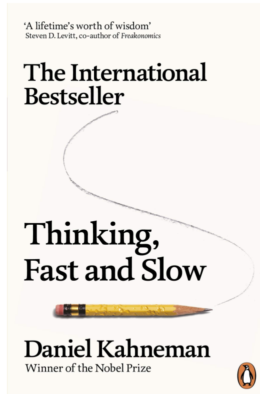

Figure 2.5: Kahneman (2011): Thinking, Fast and Slow

integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty" (The Royal Swedish Academy of Sciences 2002).

The strength of evidence from a controlled experiment is generally considered to be strong. However, the external validity, i.e., the generalizability, should be considered as well. External validity is sometimes low because effects that you can identify and measure in a lab are sometimes only of minor importance in the field.

There are different types of experiments:

Randomized controlled trials (RCTs) are a specific type of an experiment that involve randomly assigning participants to different treatment groups and comparing the outcomes of those groups. RCTs are often considered the gold standard of experimental research because they provide a high degree of control over extraneous variables and are less prone to bias.

For a better explanation and some great insights into what an RCT actually is, please watch the video produced by UNICEFInnocenti and published on the YouTube channel of UNICEF's dedicated research center https://youtu.be/Wy7qpJeozec

Quasi-experiments involve the manipulation of an independent variable, but do not involve random assignment of participants to treatment groups. Quasi-experiments are less controlled than RCTs, but can still provide valuable

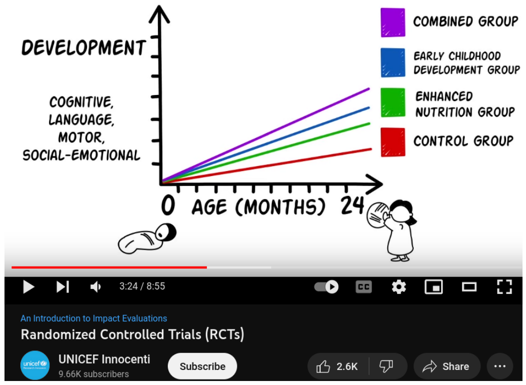

Figure 2.6: Randomized Controlled Trials (RCTs)

insights into cause-and-effect relationships.

Natural experiments involve the observation of naturally occurring events or situations that provide an opportunity to study cause-and-effect relationships. Natural experiments are often used when it is not possible or ethical to manipulate variables experimentally.

In a **laboratory experiment**, researchers manipulate an independent variable and measure the effect on a dependent variable in a controlled laboratory setting. This allows for greater control over extraneous variables, but the results may not generalize to real-world situations.

In a **field experiment**, researchers manipulate an independent variable and measure the effect on a dependent variable in a natural setting, rather than in a laboratory. This allows researchers to study real-world phenomena, but it can be more difficult to control for extraneous variables.

2.1.5 Observational data

Figure 2.7: Observational data[5]

Observational data are data that had been observed before the research question was asked or being collected independently from the study. To understand how observational data can be used to constitute a causal relationship is a bit tricky because there is only one world and only one reality at a time. In other words, we usually miss a counterfactual which we can use for a comparison. Take, for example, the past COVID-19 pandemic, where you chose to be vaccinated or not. Regardless of what you chose, we will never find out what would have happened to you if you had chosen differently. Maybe you would have died, maybe you would have gotten more or less sick, or maybe you wouldn't have gotten sick at all. We don't know, and it's impossible to find out because it's impossible to observe the counterfactual outcomes. This makes it difficult to establish causality from observational data. However, ingenious minds have found reasonable procedures and methods to extract some level of knowledge from observational data that allows us to infer causal relationships from observational data where we cannot directly observe the counterfactual outcome. We will come back to these methods later on.

In the upcoming sections, however, we will discuss experimental research designs including *randomized controlled trials* (RCTs) which are considered to be the "gold standard for measuring the effect of an action" (Taddy 2019, 128). RCTs can be used, for example, to study the effectiveness of drugs by observing people randomly assigned to three groups, one taking the pill (or treatment), a second receiving a placebo, and a third taking nothing. If the first group responds in any way differently than the other groups, the drug has an effect. Before explaining an RCT in more detail, we need to be clear about the fundamental problem of causal inference. This will be discussed in the following.

2.2 Correlation does not imply causation

Correlation refers to a statistical relationship between two variables, where one variable tends to increase or decrease as the other variable also increases or decreases. However, just because two variables are correlated does not necessarily mean that one variable causes the other. This is known as the *correlation does not imply causation* principle.

[5] The picture is free to use under the Pixabay license, see: https://pixabay.com/images/id-5029286/

For example, it may be observed that the number of storks in a particular area is correlated with the birth rate of babies in that area. However, this does not mean that the presence of storks causes an increase in the birth rate. It is possible that both the number of storks and the number of babies born are influenced by other factors, such as the overall population density or economic conditions in the area.

Therefore, it is important to carefully consider all possible explanations (confounders) for a correlation and to use empirical evidence to determine the true cause-and-effect relationship between variables.

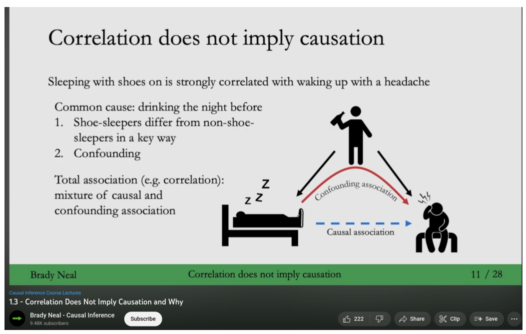

Figure 2.8: Correlation does not imply causation[6]

Watch Brady Neal's lecture on *Correlation Does Not Imply Causation and Why* which can be found here: https://youtu.be/DFPm_a-_uJM. Alternatively, you can read chapter 1.3 of his lecture notes (Neal 2020) which you find here: https://www.bradyneal.com/Introduction_to_Causal_Inference-Dec17_2020-Neal.pdf

2.3 The fundamental problem of causal inference

> Cunningham (2021, ch. 1.3): *"It is my firm belief, which I will emphasize over and over in this book, that without prior knowledge, estimated causal effects are rarely, if ever, believable. Prior knowledge is required in order to justify any claim of a causal*

[6] Picture is taken from https://youtu.be/DFPm_a-_uJM

Figure 2.9: The book cover of "Causal Inference: The Mixtape?" by Scott Cunningham (2021)

> *finding. And economic theory also highlights why causal inference is necessarily a thorny task."*

As Cunningham (2021) explains in his book, it is very hard to claim causality. In the following section, I will paraphrase briefly two aspects why it is so difficult to claim to have found a causal effect. It is difficult to know how to use data properly and how to find or generate the right data. Therefore, I will discuss Simpson's Paradox and the that gives an idea how difficult it is to analyze observational data meaningful and that we need to have a theory when looking on data and we should try to challenge the assumptions on which out theory is build on. After that I will briefly discuss

Exercises (Causal Inference ch.1)

Before going on with the content in these notes, please read chapter 1 (Introduction) of Cunningham (2021) and answer the following questions. The book is freely available (https://mixtape.scunning.com/) and the link to chapter 1 is https://mixtape.scunning.com/01-introduction.

1. What are some common misconceptions about causality that the author addresses in chapter 1?
2. What is the role of randomization in causal inference, as described in the book?

Answers:

1. Some common misconceptions about causality that the author addresses in chapter 1 include confusion between correlation and causality, and the belief that observational studies cannot (hardly) establish causality without prior knowledge. He says that human beings "engaging in optimal behavior are the main reason correlations almost never reveal

causal relationships, because rarely are human beings acting randomly" which is crucial for identifying causal effects.

2. The role of randomization in causal inference, as described in the book, is that it helps to control for confounding variables and allows for the estimation of causal effects.

2.3.1 Simpsons Paradox

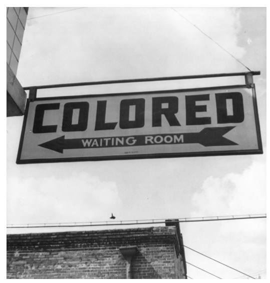

Figure 2.10: Discrimination[7]

Discrimination is bad. Whenever we see it, we should try to find ways to overcome it. De jure segregation mandated the separation of races by law is clearly discriminatory. Other forms of discrimination, however, are often more difficult to spot and as long we don't have good evidence for discrimination, we should not judge prematurely. That means we should be sure that we see an act of making unjustified distinctions between individuals based on some categories to which they belong or perceived to belong. For example, if men and women are treated differently without an acceptable reason, we consider it discriminative. For example, UC Berkeley was accused of discrimination in 1973 because it admitted only 35% of female applicants but 44% of male applicants overall. The difference was statistical significant. However, it turned out that the selection of students was not discriminative against women but agains men accordingly to Bickel, Hammel, and O'Connell (1975). Who conclude there was just a "tendency of women to apply to graduate departments that are more difficult for applicants of either sex to enter" (Bickel, Hammel, and O'Connell 1975, 403). Figure 2.11 taken from Bickel, Hammel, and O'Connell (1975, 403) visualizes this fact.

[7] The photography is public domain and stems from the Library of Congress Prints and Photographs Division Washington, see: http://hdl.loc.gov/loc.pnp/pp.print.

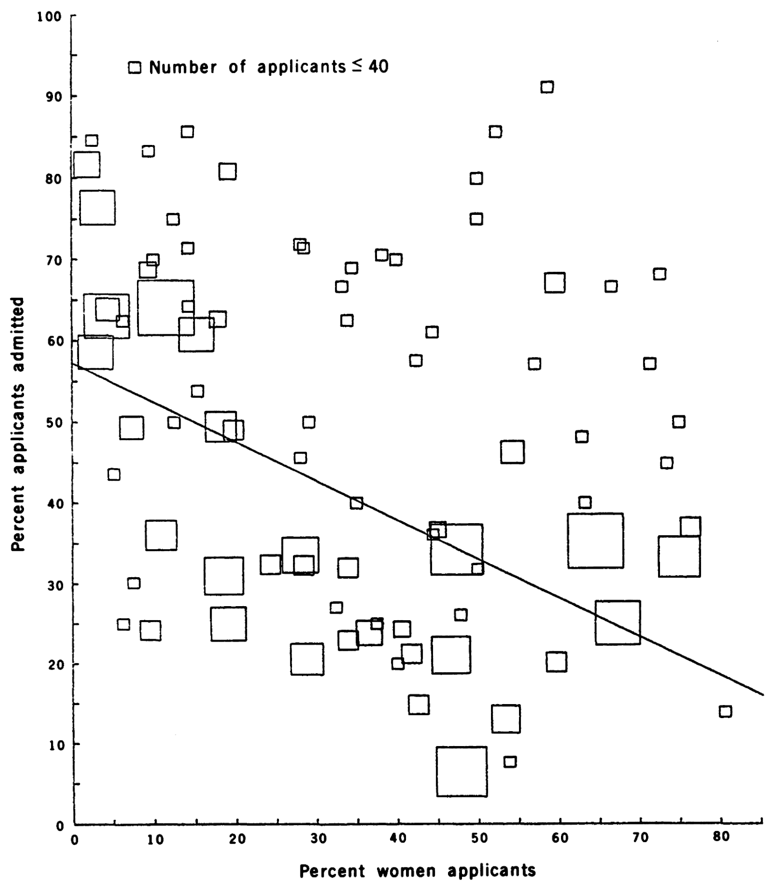

Fig. 1. Proportion of applicants that are women plotted against proportion of applicants admitted, in 85 departments. Size of box indicates relative number of applicants to the department.

Figure 2.11: Graph taken from Bickel, Hammel, and O'Connell (1975, 403)

Here you can read the summary of their infamous study:

> "Examination of aggregate data on graduate admissions to the University of California, Berkeley, for fall 1973 shows a clear but misleading pattern of bias against female applicants. Examination of the disaggregated data reveals few decision-making units that show statistically significant departures from expected frequencies of female admissions, and about as many units appear to favor women as to favor men. If the data are properly pooled, taking into account the autonomy of departmental decision making, thus correcting for the tendency of women to apply to graduate departments that are more difficult for applicants of either sex to enter, there is a small but statistically significant bias in favor of women. The graduate departments that are easier to enter tend to be those that require more mathematics in the undergraduate preparatory curriculum. The bias in the aggregated data stems not from any pattern of discrimination on the part of admissions committees, which seem quite fair on the whole, but apparently from prior screening at earlier levels of the educational system. Women are shunted by their socialization and education toward fields of graduate study that are generally more crowded, less productive of completed degrees, and less well funded, and that frequently offer poorer professional employment prospects."

Exercises (graduate admissions)

Read the first three pages of Bickel, Hammel, and O'Connell (1975), i.e., pages 398-400, and answer the following questions. The article can be found here: https://www.science.org/doi/pdf/10.1126/science.187.4175.398.

a) Describe the two assumptions that must be true in order to prove that UC Berkeley discriminates against women or men overall.
b) Table 1, shows that 277 fewer women and 277 more men were admitted than we would have expected under the two assumptions. Show how this number was calculated.
c) Explain the analogy with fish that illustrates the danger of pooling data.

Answers:

a) Assumption 1 is that in any given discipline male and female applicants do not differ in respect of their intelligence, skill, qualifications, promise, or other attribute deemed legitimately pertinent to their acceptance as students. It is precisely this assumption that makes the study of "sex bias" meaningful, for if we did not hold it any differences in acceptance of applicants by sex could be attributed to differences in their qualifications, promise as scholars, and so on. (...) Assumption 2 is that the sex ratios of applicants to the various fields of graduate study are not importantly associated with any other factors in admission." (Bickel, Hammel, and O'Connell 1975, 398)
b) Expectations were taken based on the overall acceptance rate of about 0.416 and multiplied by the total observed numbers of applicants admitted and rejected. For example: $(3738 + 4704) \cdot 0.41666 \approx 3460$ and $(3738 + 4704) \cdot (1 - 0.41666) \approx 4981$. Taking the difference of these two measures gives the number to be explained.
c) The analogy is explained on page 400: "Picture a fishnet with two different mesh sizes. A school of fish, all of identical size (assumption 1), swim toward the net and seek to pass. The female fish all try to get through the small mesh, while the male fish all try to get through the large mesh. On the other side of the net all the fish are male. Assumption 2 said that the sex of the fish had no relation to the size of the mesh they tried to get through. It is false."

The UC Berkley case is just one of many examples to illustrate that uniformity of group assignment of individuals is a necessary condition to ensure that pooling of data does not lead to misleading conclusions when using statistics. The phenomenon of obtaining different results depending on whether one considers the data pooled or unpooled is often referred to as the *Simpson paradox*.

Exercises (Simpson's Paradox)

1. What is Simpson's Paradox?
   a) A phenomenon in which the direction of a relationship between two variables changes when a third variable is introduced
   b) A phenomenon in which the strength of a relationship between two variables changes when a third variable is introduced
   c) The phenomenon where correlation appears to be present in dif-

ferent groups of data, but disappears or reverses when the groups are combined

2. What is a potential cause of Simpson's Paradox?
   a) Differences in the variance of the two variables
   b) Differences in the correlation of the two variables
   c) Confounding variables
   d) Differences in the sample size of the two variables

Answers: 1. a), 2. c) and d)

2.3.2 Rubin causal model

> Keele (2015, 314): *"An identification analysis identifies the assumptions needed for statistical estimates to be given a causal interpretation."*

If we are interested in the causal effect of a certain treatment on an outcome, we need to compare the outcome of the individuals who received the treatment to the outcome of the individuals who did not receive the treatment. However, if the counterfactual outcome is missing for some individuals, we cannot make this comparison and therefore cannot estimate the causal effect. Unfortunately, the counterfactual is very often missing. For example, if we want to measure the effect of a vaccine we never can have a person who is vaccinated and not vaccinated at the same time.

The Rubin Causal Model, also known as the potential outcomes framework, is a statistical framework for analyzing causality in the context of missing data. It goes back to Donald B. Rubin (*1943) a statistician and is now a widely used method for causal inference. The basic premise of the Rubin Causal Model is that for each individual in a study, there are two potential outcomes: the outcome that would occur if the individual were exposed to a certain treatment or intervention (the "treatment group"), and the outcome that would occur if the individual were not exposed to that treatment (the "control group"). The key idea is that these potential outcomes can be used to infer causality by comparing the outcomes between the treatment and control groups even if we do not have a full set of data.

Watch Brady Neal's lecture on *What Does Imply Causation? Randomized Con-*

[8] Picture is taken from https://youtu.be/gGaWU8XEoGk

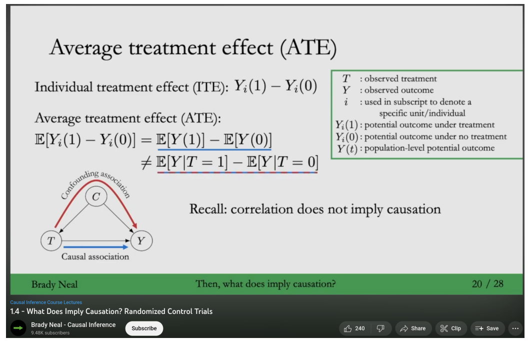

Figure 2.12: Average treatment effect (ATE)[8]

trol Trials which can be found here: https://youtu.be/gGaWU8XEoGk. Alternatively, you can read chapter 2 of his lecture notes (Neal 2020) which you find here: https://www.bradyneal.com/Introduction_to_Causal_Inference-Dec17_2020-Neal.pdf

Given the some assumptions, the Rubin Causal Model allows for estimation of the average treatment effect (ATE) - the difference in the expected outcomes between the treatment and control groups. There are several ways to estimate the ATE under the Rubin Causal Model some of which will be part of this course. Applied correctly, it can provide valuable insights into causality and inform decision making. However, the Rubin Causal Model has its limitations and assumptions that need to be met in order for the inferences to be valid.

Exercises (treatment effects)

1. What is the individual treatment effect (ITE)?
2. What is the average treatment effect (ATE)?
3. How is the ATE calculated?
4. Can the ATE be used to determine the effect of a treatment on an individual level?
5. What are some potential sources of bias when estimating the ATE?

Answers:

1. The individual treatment effect (ITE) is a measure of the effect of a treatment or intervention on an individual level. It represents the difference in the outcome for an individual who receives the treatment

versus the outcome for that same individual if they had not received the treatment.

2. The average treatment effect (ATE) is a measure of the difference in the expected outcomes between a treatment group and a control group. It represents the overall effect of a treatment on the population as a whole.
3. The ATE is calculated by taking the difference between the average outcome for the treatment group and the average outcome for the control group.
4. No, the ATE is a population-level measure and cannot be used to determine the effect of a treatment on an individual level. To determine the effect of a treatment on an individual level, you would need to use techniques such as propensity score matching or instrumental variables.
5. Some potential sources of bias when estimating the ATE include selection bias, measurement bias, and unobserved confounding variables. To mitigate these biases, researchers may use randomization or other advanced statistical techniques such as propensity score matching or instrumental variables to control for these potential sources of bias.

Chapter 3

Difference in difference

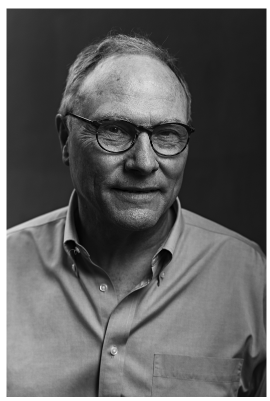

Figure 3.1: David Card (*1956)[1]

David Card is one of the most influential labor economist of the 20th century and Nobel laureate of 2021. He is well-known for his research on the effects of the minimum wage on employment, which challenged the traditional view that increasing the minimum wage leads to a decrease in employment. In his article *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania* (Card and Krueger 1994) he and Alan Krueger (1960-2019) used a *natural experiment* to examine the effect of an increase in the minimum wage on employment. In particular, they identified a treatment group (restaurants in New Jersey) and a control group

[1]Photo stems from his public homepage https://davidcard.berkeley.edu/

(restaurants in eastern Pennsylvania) to measure the effect of increasing the minimum wage that was increased in New Jersey but not in Pennsylvania. This increase did not lead to a decrease in employment, which contradicted the widely held view that increasing the minimum wage would lead to job loss. The empirical method that they used is called *difference in difference* and we discuss it in the following section.

The difference in difference (DiD) method allows to estimate the causal effect of a treatment or intervention. In particular, it is popular to study the impact of policy changes and other interventions on a specific outcome of interest.

The basic idea behind the DiD method is to compare the change in an outcome variable between a treatment group and a control group over time. The treatment group is the group that is exposed to the intervention or treatment, while the control group is a group that is not exposed to the intervention. The difference in the change in the outcome variable between the two groups is then used to estimate the causal effect of the intervention.

To use the DiD method, researchers typically collect data on the outcome variable of interest for both the treatment and control groups before and after the intervention. This data is then used to calculate the difference in the change in the outcome variable between the two groups.

For example, if a study aims to examine the effect of a new policy on the employment rate, it should collect data on the employment rate for a group of individuals living in a region where the policy was implemented, and for a group living in a similar region where the policy was not implemented. The study can then compare the change in the employment rate for the two groups, before and after the implementation of the policy. The difference in the change in the employment rate between the two groups can be used to estimate of the causal effect of the policy on employment.

It is important to note that DiD assumes that there are no other factors that could be affecting the outcome variable of interest and that the treatment and control groups are similar in all ways except for the intervention. To control for these assumptions researchers can use statistical techniques such as matching to ensure that treatment and control groups are similar before the intervention.

DiD is useful when we only have observational data and in situations where it is not possible or ethical to randomly assign individuals to a treatment or control group, for example, in the case of policy changes.

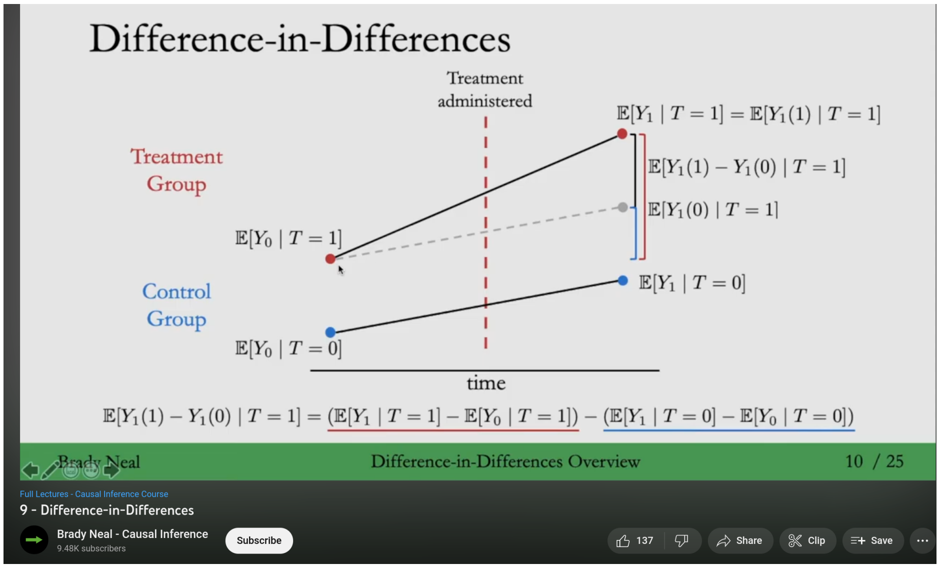

Figure 3.2: Differences-in-Differences

The figures above stem from Brady Neal's lecture on *Difference-in-Differences* which can be found here: https://youtu.be/2nDgrNP7XSE. Please watch this video.

Exercises (DiD)

Watch the video *Introduction to Differences-in-Differences* found here: https://youtu.be/eiffOVbYvNc and answer the following questions. The video is part of a course called *Mastering Econometrics with Joshua Angrist (MIT)* produced by Marginal Revolution University. In it, Josh Angrist introduces differences within differences using one of the worst economic events in history: the Great Depression.

1. In the video, the treatment being examined is:

a) Bank failure.
b) "Easy" money.
c) "Tight" money.

[3]The picture stems from the video https://youtu.be/eiffOVbYvNc

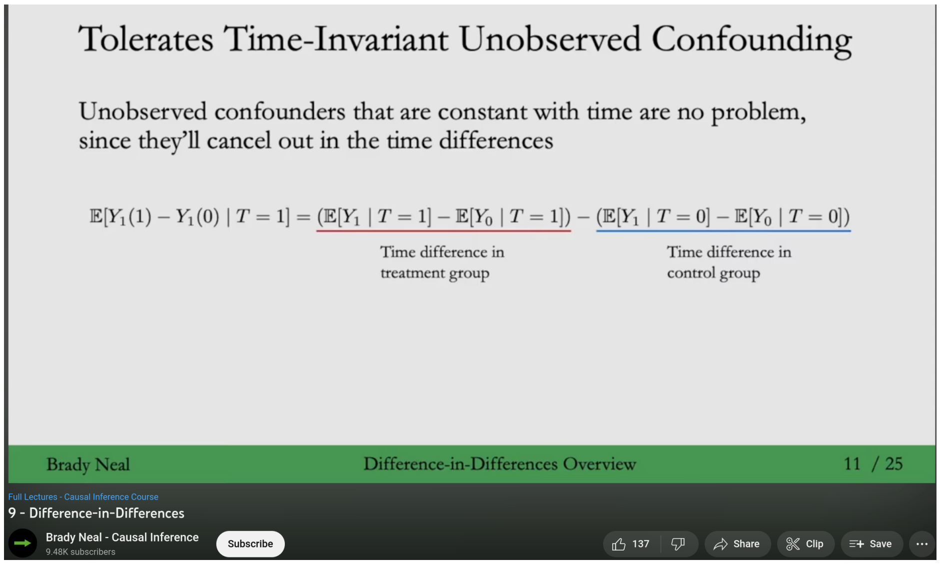

Figure 3.3: Tolerate time-invariant unobserved confounding

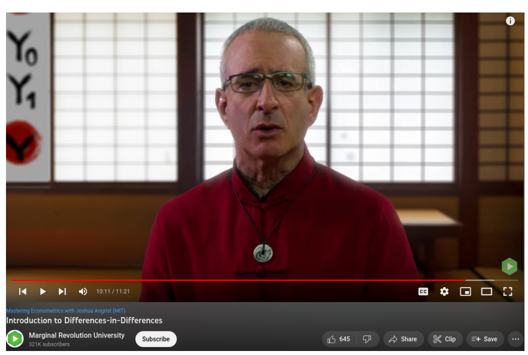

Figure 3.4: Josh Angrist (*1960): Nobel Prize winner of economics in 2021[3]

d) Differences-in-Differences.
e) None of the above.

2. If the treatment were effective, which outcome would we expect to observe?

a) Fewer bank failures.
b) Increased bank failures.
c) Continued parallel trends.
d) No differences in any variables unrelated to bank failure.
e) None of the above.

3. Practically, how is DD (Differences-in-Differences) typically implemented?

a. Non-parametric statistical techniques.
b. Randomized trials.
c. Regression analysis.
d. Instrumental variables.
e. None of the above.

Answers: 1. b), 2. a), 3. c)

Literature

Bickel, Peter J, Eugene A Hammel, and J William O'Connell. 1975. "Sex Bias in Graduate Admissions: Data from Berkeley: Measuring Bias Is Harder Than Is Usually Assumed, and the Evidence Is Sometimes Contrary to Expectation." *Science* 187 (4175): 398–404. https://doi.org/10.1126/science.187.4175.398.

Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *The American Economic Review* 84 (4): 772–93.

Cunningham, Scott. 2021. *Causal Inference*. Yale University Press. https://mixtape.scunning.com.

Huntington-Klein, Nick. 2022. *The Effect: An Introduction to Research Design and Causality*. CRC Press. https://theeffectbook.net.

Hurston, Zora Neale. 2010. *Dust Tracks on a Road*. Harper Perennial Modern Classics. New York, NY: HarperCollins.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Penguin.

Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23 (3): 313–35.

Neal, Brady. 2020. "Introduction to Causal Inference from a Machine Learning Perspective." Online. https://www.bradyneal.com/Introduction/_to/_Causal/_Inference-Dec17/_2020-Neal.pdf.

Taddy, Matt. 2019. *Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions*. 1st ed. McGraw Hill Education.

The Royal Swedish Academy of Sciences. 2002. Press Release. https://www.nobelprize.org/prizes/economic-sciences/2002/press-release/.

Weymar, Paul. 1955. *Konrad Adenauer: Die Autorisierte Biographie*. Kindler.