Experimental economics

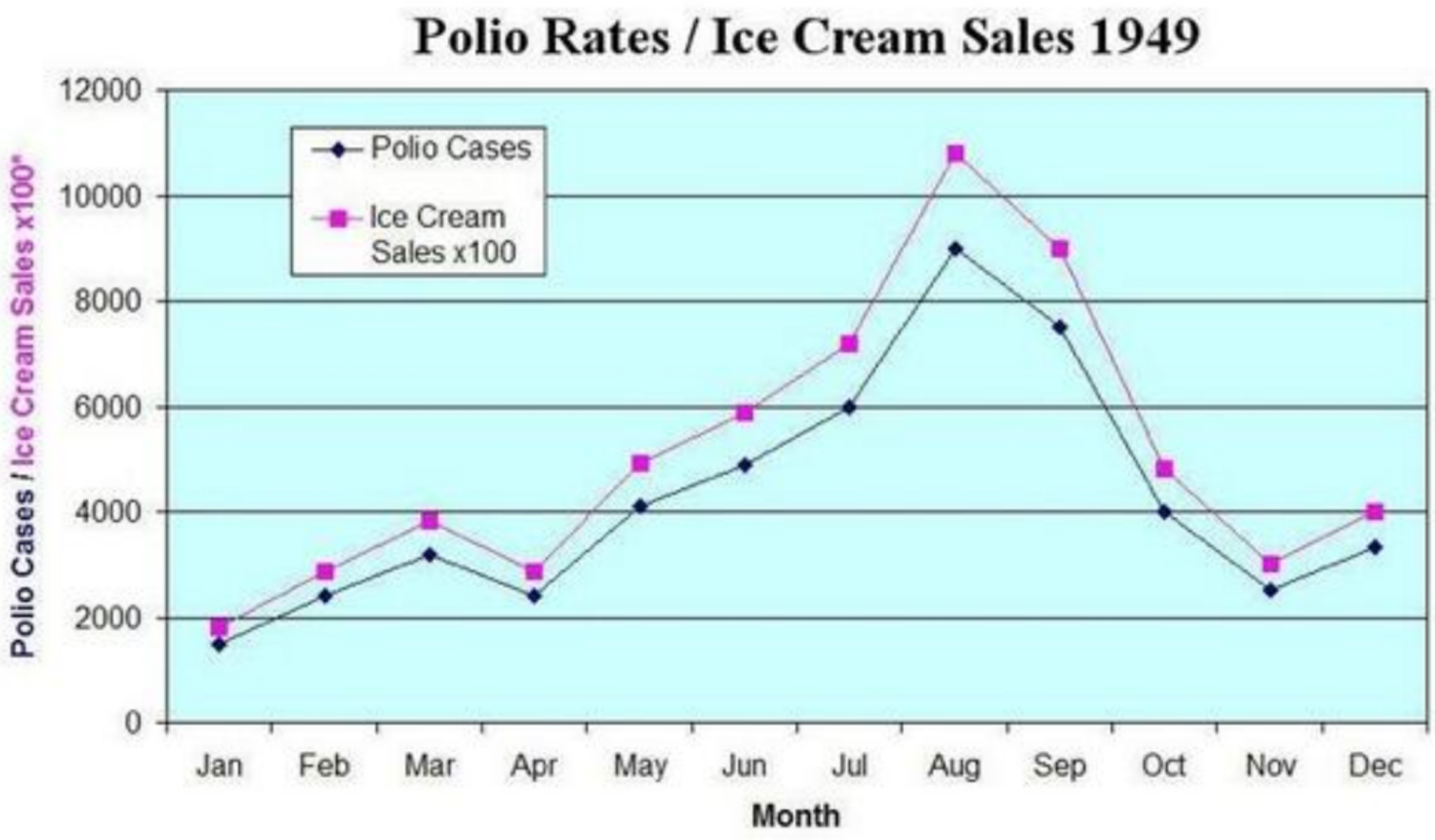
Lecture 1: Economics as an experimental science

Matej Lorko
matej.lorko@euba.sk
Materials: www.lorko.sk/lectures

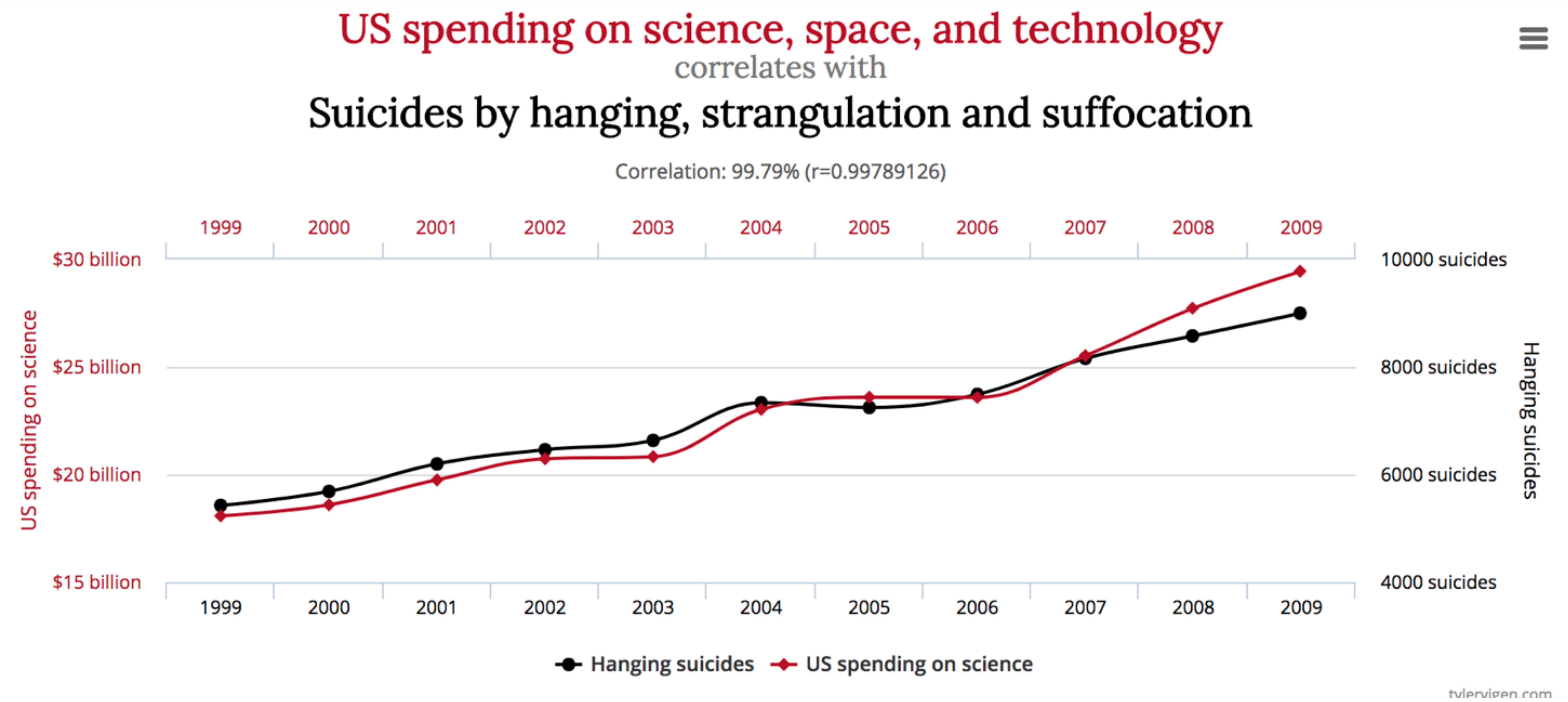
References:

- Jacquemet, N., & l'Haridon, O. (2018). Experimental economics. Cambridge University Press.
- Weimann, J., & Brosig-Koch, J. (2019). Methods in experimental economics. Springer International Publishing. Chicago

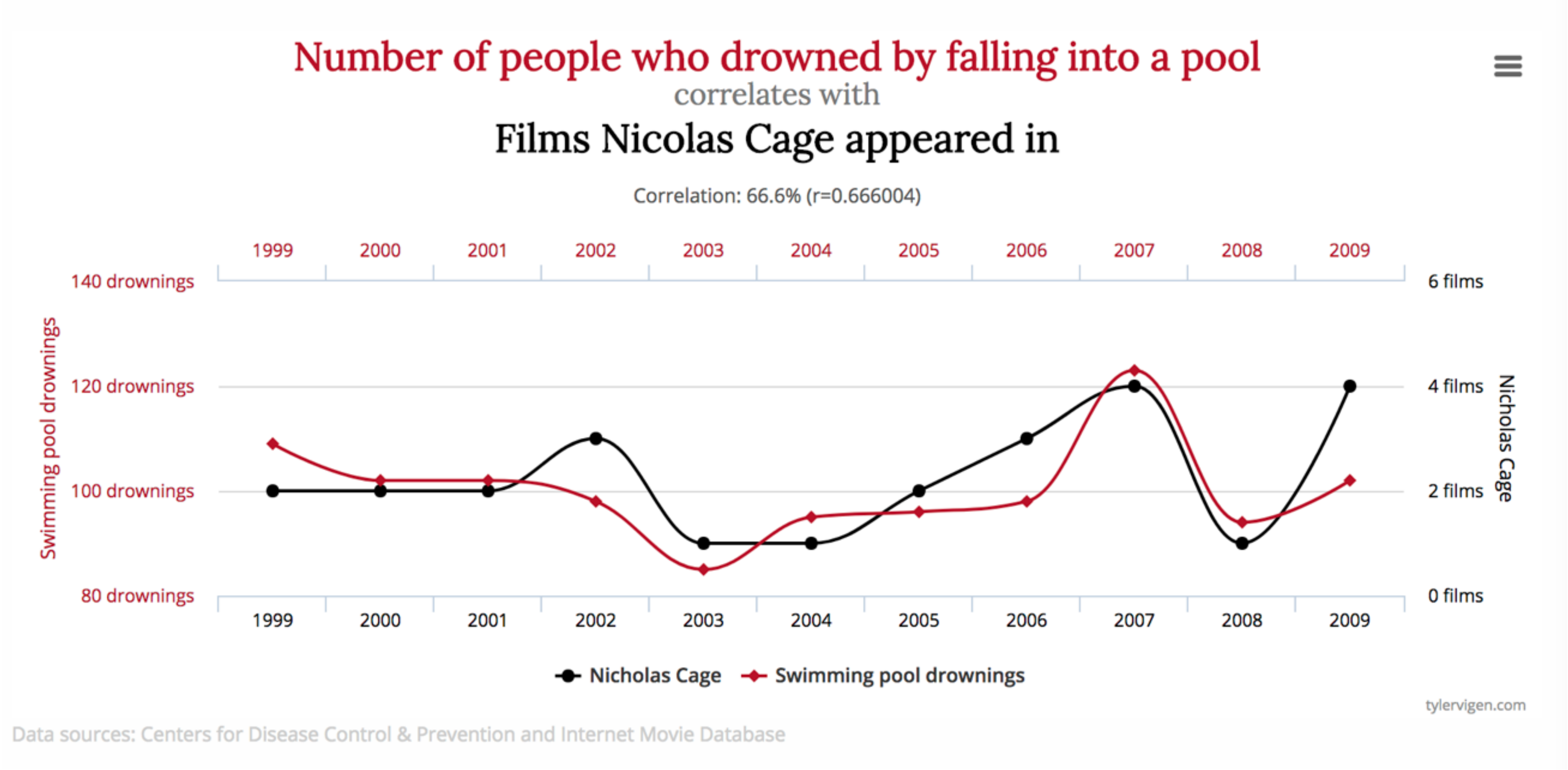
The Real Cause of Polio!



Correlation vs. Causation



Correlation vs. Causation




Correlation
Causality

Correlation vs. causation

- Research in labour economics shows that higher performance is usually observed in firms paying a piece rate rather than a fixed wage. But it is so because piece rates induce higher performance, or just the reverse: that firms with a piece rate compensation scheme attract higher-performance workers?
- Similarly, the likelihood of death and the time spent at hospital are strongly positively related in any population; do such data inform us about how dangerous hospitals are for health?
- The number of policemen in a geographic area is often positively correlated to crime rates; does it mean one should reduce police forces to contain crime?
- Unemployed people who receive more help from public placement agencies generally experience lower likelihood of finding a job; do placement agencies hurt the labour market potential of job-seekers?
- More examples:
    - Cholera in Russia
    - SAT preparation courses vs. test scores
    - Breast-feeding vs. malnutrition
    - U20 ice hockey rosters

Correlation vs. causation

- Historically, economics has relied on observation of naturally occurring processes for its data needs
  - Advantage: these processes are observed in their undisturbed form
  - Disadvantage: it is diffucult to establish causal relationships among observed variables since many of them are determined simultaneously
- Fundamental issue faced by those doing empirical work in economics: disentangling causality from correlation.We say that two economic variables are correlated if they move together. But this relationship is causal only if one of the variables is causing the movement in the other. The general problem that empirical economists face in trying to use existing data to assess the causal influence of one factor on another is that one cannot immediately go from correlation to causation. Knowing that two factors are correlated provides no predictive power; prediction requires understanding the causal links between the factors.
- Similar problems appear in public policies. For example:
  - Do lower welfare benefits cause higher labor supply among single mothers? Do larger benefits for unemployment insurance cause individuals to stay unemployed longer? How does minimum wage affect the employment?
  - Do education raise wages? Does class size affect learning?
  - Does access to information improve the market effectiveness?
- Identification problem: given that two series are correlated, how do you identify whether one series is causing another?

Correlation vs. causation

- Analysis
  - Step 1: Document the correlation, that is whether data on two measures move together.
  - Step 2: Assess whether the movements in one measure are causing the movements in the other.
    - For any correlation between two variables A and B, there are three possible explanations, one or more of which could result in the correlation:
      - A is causing B.
      - B is causing A.
      - Some third factor is causing both.
  - Well designed research thus needs to make valid causal inferences. Ideally, such a design does three things:
    - 1. Covariation: demonstrates that the alleged cause (call it X) does in fact covary (corelate) with the supposed effect, Y.
    - 2. Time order: The research must show that the cause preceded the effect: X must come before Y in time. After all, can an effect appear before its cause?
    - 3. Elimination of possible alternative causes, sometimes termed "confounding factors": The research must be conducted in such a way that all possible joint causes of X and Y have been eliminated.

Assessing causation

- SAT:
  - A -> B: SAT prep courses worsen preparation for SATs.
  - B -> A: Those who are of lower test-taking ability take preparation courses to try to catch up.
  - C -> A,B: Those who are generally nervous people like to take prep courses, and being nervous is associated with doing worse on standardized exams.
- Breast-feeding:
  - A -> B: Longer breast-feeding is bad for health.
  - B -> A: Those infants who are in the worst health get breast-fed the longest.
  - C -> A,B: The lowest-income mothers breast-feed longer, since this is the cheapest form of nutrition for children, and low income is associated with poor infant health.

The problem of bias

- We should always start our analysis of an empirical methodology with a simple question: Do the treatment and control groups differ for any reason other than the treatment?

- The non-treatment-related differences between treatment and control groups are the fundamental problem in assigning causal interpretations to correlations. We call these differences **bias**, a term that represents any source of difference between treatment and control groups that is correlated with the treatment but is not due to the treatment.

- By definition, such differences do not exist in a randomized trial, since the groups do not differ in any consistent fashion, but rather only by the flip of a coin.

- Thus, randomized treatment and control groups cannot have consistent differences that are correlated with treatment, since there are no consistent differences across the groups other than the treatment. As a result, randomized trials have no bias, and it is for this reason that randomized trials are the gold standard for empirically estimating causal effects.

- The description of randomized trials here relies on those trials having fairly large numbers of treatments and controls (large sample sizes). Having large sample sizes allows researchers to eliminate any consistent differences between the groups by relying on the statistical principle called the **law of large numbers**: the odds of getting the wrong answer approaches zero as the sample size grows.

Golden standard for assessing causation: Randomized trials

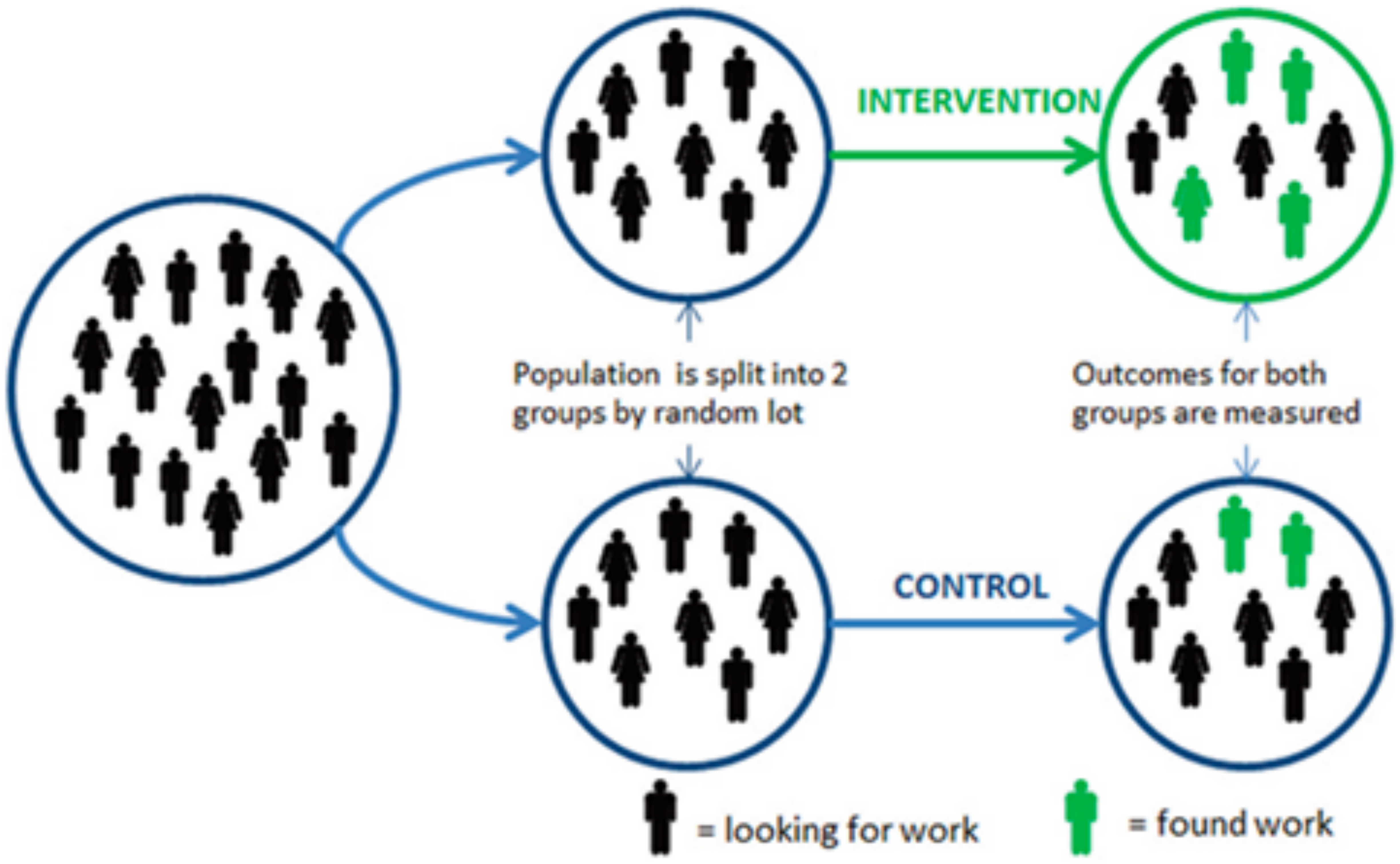


Figure 1. The basic design of a randomised controlled trial (RCT), illustrated with a test of a new 'back to work' programme.

Experiments in economics

- “There is a property common to almost all the moral sciences, and by which they are distinguished from many of the physical; that is, that it is seldom in our power to make experiments in them.” - Mill (1836), cited in Guala (2005, p. 2).
- “Economists (unfortunately) cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors. Like astronomers or meteorologists, they generally must be content largely to observe.” - Samuelson and Nordhaus (1985), Principles of Economics.
- “As currently practiced, economics is ideally suited for experimental investigation.” - John Hey (1991), Experiments in Economics
- “Experimental economics is an exciting new development.” - (Samuelson and Nordhaus, 1992, p. 5)
- Nobel Prizes
    - Kahneman (2002) for having integrated insights from psychological research into economic science, especially concerning human judgement and decision-making under uncertainty
    - Smith (2002) for having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms
    - Roth (2012) for the theory of stable allocations and the practice of market design
    - Thaler (2017) for his contributions to behavioral economics
    - Banerjee, Duflo, Kremer (2019) for their experimental approach to alleviating global poverty

Experimental economics

- Economics is a “behavioral science” since it focuses on people’s decisions on the use of scarce resources. Since about the beginning of the twentieth century economics has stood apart from other disciplines dealing with human behavior in its broadest sense due to the fact that it uses formal mathematical models and abstract theories based on clearly defined assumptions.

- Economic research based on models allows the assumptions on which a scientific statement is based to be precisely stated in a mathematical sense, thus enabling us to very precisely specify the situation to which the theory is to be applied. The decisive factor here is that the formal method allows the conclusions that economists draw to be derived and proven from the assumptions with exactly the same precision. Economists can therefore make “if-then” statements whose clarity and precision would be unimaginable without formal methods.

- At first sight, economists pay a high price for the clarity of these statements. A real-world test of the theory is scarcely possible because the assumptions used in the models economists construct have a high degree of abstraction from the real conditions of the everyday economic world and because they use ideal-typical behavioral models. And what good is a theory that makes unambiguous and mathematically elegant statements if it is not possible to know whether these statements have any significance for economic reality? This “empirical weakness” is transformed into a strength when the possibility of verifying the theory experimentally.

- It is no longer possible to imagine the economic sciences without experimental research. It has become a well-established method and is used in virtually all branches of economics. Today it is a matter of course that experimental work is published in nearly all international economic journals and are regularly featured in the top journals. This has not always been the case. Only 30 years ago, experimental work was an absolute rarity in important journals and experimental research on a wider scale has only existed since the 1970s.

Experimental economics vs. behavioral economics

- Behavioral economics: research programme aimed at improving economic analysis using realistic psychological assumptions about human behavior.

- Experimental economics is not a reseach programme. Reather, it is a research method based on experimental control, applied to the typical topics in economic analysis. However, it is particularly suited for the study of the phenomena of interest to the behavioral economics.

- Both behavioral and experimental economics owe a great deal to the accumulated knowledge in experimental psychology. Controlled experiments have been used for a long time in this field and most methodological discussions took place before they even appeared in economics.

Pioneering experiments

- As early as 1931 Louis Leon Thurstone attempted to derive indifference curves experimentally by offering subjects a (hypothetical) choice between alternative bundles of goods. Thurstone himself was not an economist, but a psychologist, and the experimental methods he used were more consistent with those of experimental psychology and less with those used by experimental economists today. The criticism that directed at this experiment said that essentially nothing can be learned from hypothetical questions because the subjects are not provided with the right incentives.

- Of a somewhat different nature is the experiment published in 1948 by Edward Hastings Chamberlin. He was the first to attempt to create markets in the laboratory, applying methods that are still also used today in an enhanced form to conduct market experiments. Further development of his method by Vernon Smith later led to a convincing confirmation in the laboratory that even under difficult conditions (e.g., in the case of very limited information) markets are indeed able to generate equilibrium prices.

- In 1950 Melvin Dresher and Merrill Flood devised a game based game-theoretical analysis on that has since then had a remarkable career in economics: the prisoner's dilemma. These two authors thought up this game with the intention of subjecting the concept of the Nash equilibrium to a particularly hard test (today this would be called a stress test).

- The closeness of the relationship between game theory and experimental research is also evident in the fact that a number of outstanding game theoreticians are among the early experimentalists. Expected utility theory and, as a consequence, also game theory use the notion of optimizing players who act strictly rationally as a basic premise. Beginning with the early experiments on the prisoner's dilemma of Dresher and Flood and the experiments that tested expected utility theory, the history of experimental research has also time and again been a history of findings that are at odds with the assumption of rationality. This does not mean that experiments always show non-rational behavior, but it occurs relatively frequently.

Experiments in economics

- It is mainly thanks to two important characteristics that experiments are so well suited to testing formal theory. First, experiments make it possible to vary the conditions under which decisions are made in a targeted and controlled way. The experimenter is in control of what information he provides the subject, and can thus systematically investigate what influence the extent to which the decision-maker is informed has on the decision.

- This possibility to adjust the variables that are important for behavior in a controlled way represents a second very important feature of experiments. It allows the researcher to pose the question that specifically interests him and to gather the data that is specifically relevant to this question. He is not reliant on economic reality to provide the data he requires to investigate a particular issue. Rather, he is in a position to generate the data to virtually every question that can be asked.

- This means, though, that the opportunities offered by experimental research go beyond merely verifying theories. It can also be used to search for stylized facts, regularities in behavior that have not, at least not yet, been described by theories. By this means, explorative experiments yield observations that could provide valuable information as to how successful behavioral theories can be descriptively formulated.

Experiments in economics

- The range of application of the experimental method is therefore not limited to those areas determined by existing theories. Roth (1995) once aptly described this by attributing three main functions to experiments. They can:
    - “speak to theoreticians” by testing theories, helping to find new theories and establishing causal relationships
    - “search for facts” by uncovering stylized facts, for example investigating gender differences in attitudes toward risk and competition, or comparing efficiency of various market institutions
    - “whisper in the ear of princes”, i.e. they can be used to provide policy advice

Testing theories

- Theoretical models rely on behavioural assumptions to provide an understanding of the decisions of agents, and the resulting outcomes, induced by a given environment. They do so by restricting the economically relevant situation to a few key features.

- For example, auction models reduce the auction environment to marginal values for the good, prices and monetary benefits. Based on the assumptions of utility maximisation and the axiomatic underlying the game-theoretic analysis of strategic interactions, it yields clear-cut predictions of both bidding behaviour and the properties of the resulting allocation.

- Experiments exhibit two major advantages in that regard: the ability to both build an empirical situation that mimics the theoretical model, and measure or observe usually non-observable, or hardly measurable, dimensions (such as individual preferences towards the good, or individual prices posted). Experiments also get theory closer to reality by providing measures of individual preferences.

- Experiments not only allow us to assess whether the theoretical account of preferences actually makes empirical sense (e.g. to what extent is behaviour in risky environments actually described by the assumptions of expected utility), but also to assess whether predicted behaviour based on such preferences coincides with what theory predicts. In all these instances, experiments help in assessing the empirical relevance of theoretical results in terms of accuracy, precision and extent.

Searching for facts

- There are many economic situations that are worth understanding, but which are too complicated and/or too specific to be covered by theory.

- Auctions again provide a useful illustration of such experiments: as the allocation mechanism, or the amount of information available to bidders, becomes more specific, auction models quickly become intractable. This does not mean, of course, that such specific auction markets are of no economic interest (even when there is no obvious reason why a market works as it does, the mere fact that it is used in practice is sometimes enough to make it worth investigating).

- In such cases, experiments can be used as a substitute for theoretical analysis. They are used to 'search for facts' in the sense that they allow us to mimic well-defined situations and measure behaviour as well as the outcomes they generate.

- In the absence of prior expectations based on theory, such observations provide empirical knowledge about how the environment works, and what are its most sensitive features. To serve this purpose, such evidence must be robust and conclusive enough to actually serve as a stylised fact. In that regard, the replicability of experimental data, and the possibility to assess the robustness of the results through variations of the environment, are important advantages of experiments.

Whispering in the ears of princes

- This third purpose of experiments amounts to improving the decision-making process by informing regulators or decision-makers (Roth's 'princes') of the likely outcomes of new or existing public policies. The general principle is to use experiments to test-bed decision environments such as market mechanisms, policy changes or new organisational structures.

- Observed outcomes in the experiment provide insights into the likely changes in behaviour and economic outcomes raised by innovative decision environments. As such, this aim builds on the ability of experiments to both test theory and search for facts – depending on whether theoretical insights are available on the policy-relevant question under investigation.

- The specific contribution of experiments to policy design comes from the ability to answer the specific needs of decision-makers. Because all the parameters of the decision environment can be freely set in the laboratory, an experiment makes it possible to fully replicate the specific features of a given policy.

- This ability to fine-tune the experimental environment according to the requirement of the policy-relevant questions stands in sharp contrast with observational data. As compared to field experiments, laboratory experiments are cheap and easily implementable. An additional contribution of the use of experiments to policy design is their use in an instructive function. Laboratory experiments make economic reasoning more intuitive and appealing for non-academics. Even without producing any new knowledge, they can be used to make a convincing case of what the consequences will be of an intended change in the environment.

But what if randomized trials are not available?

- For many questions of interest, randomized trials are unfortunately not available, because they can be enormously expensive, take a very long time to plan and execute, and often raise difficult ethical issues (e.g new medical procedures). Moreover, even the gold standard of randomized trials has some potential problems. First, the results are only valid for the sample of individuals who volunteer to be either treatments or controls, and this sample may be different from the population at large. For example, those in a randomized trial sample may be less averse to risk or they may be more desperately ill. Thus, the answer we obtain from a randomized trial, while correct for this sample, may not be valid for the average person in the population.

- A second problem with randomized trials is that of attrition: individuals may leave the experiment before it is complete. This is not a problem if individuals leave randomly, since the sample will remain random. Suppose, however, that the experiment has positive effects on half the treatment group and negative effects on the other half, and that as a result the half with negative effects leaves the experiment before it is done. If we focus only on the remaining half, we would wrongly conclude that the treatment has overall positive impacts.

- attrition: reduction in the size of samples over time, which, if not random, can lead to biased estimates.

- If data from randomized trials are not available, we need to work with observational data = data generated by individual behavior observed in the real world, not in the context of deliberately designed experiments. For example, instead of information on a randomized trial of a new medicine, we may simply have data on who took the medicine and what their outcomes were.

- There are several well-developed methods that can be used by analysts to address the problem of bias with observational data, and these tools can often closely approximate the gold standard of randomized trials. In other words, we can use observational data to estimate causal effects instead of just correlations. The major concern is how to overcome any potential bias so that we can measure the causal relationship (if there is one).

Quasi-Experiments

- If randomized control trial is not an option, there may still be a middle-ground between RCT and correlational analyses (e.g., OLS regressions with control variables): the quasi-experiment

- the quasi-experiment is a situation that arises naturally when changes in the economic environment (such as a policy change) create nearly identical treatment and control groups that can be used to study the effect of that policy change. In a quasi-experiment, outside forces (such as those instituting the policy change) do the randomization for us.

- With quasi-experimental studies, unlike true experiments, we can never be completely certain that we have purged all bias from the treatment–control comparison. Quasi-experimental studies use two approaches to try to make the argument that they have obtained a causal estimate. The first is intuitive: trying to argue that, given the treatment and control groups, it seems very likely that bias has been removed. The second is statistical: to continue to use alternative or additional control groups to confirm that the bias has been removed.

Quasi-Experiments

- Difference-in-difference estimator: The technique that tries to combine time series and cross-sectional analyses to address the problems with each. By comparing the change in population A to the change in population B, the estimator controls for other time series factors that bias the time series analysis within population A. Likewise, by comparing the change within each population, rather than just comparing the two populations at a point in time, the estimator controls for omitted factors that bias cross-sectional analysis across the two populations.

- Searching for a change in variable X
  - 2 periods (Y, Z)
  - 2 populations (A,B)
  - In period Y, the policy is the same for A and B
  - In period Z, there is new policy for A, while the policy for B is not changed
- x (population A, year Y) - x (population A, year Z) = Treatment effect + Bias
- x (population B, year Y) - x (population B, year Z) = Bias
- Difference = Treatment effect

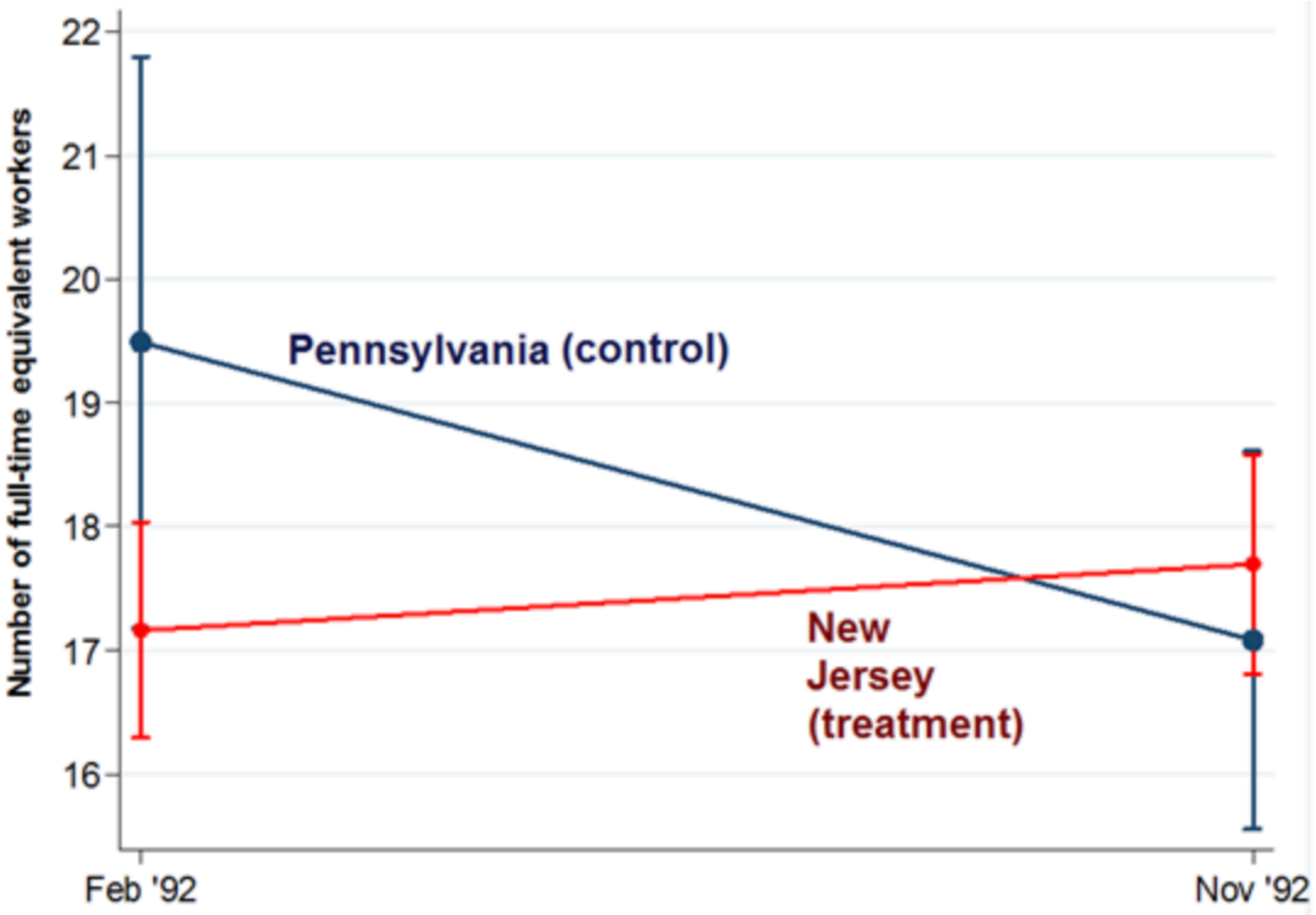
■ TABLE 3-1

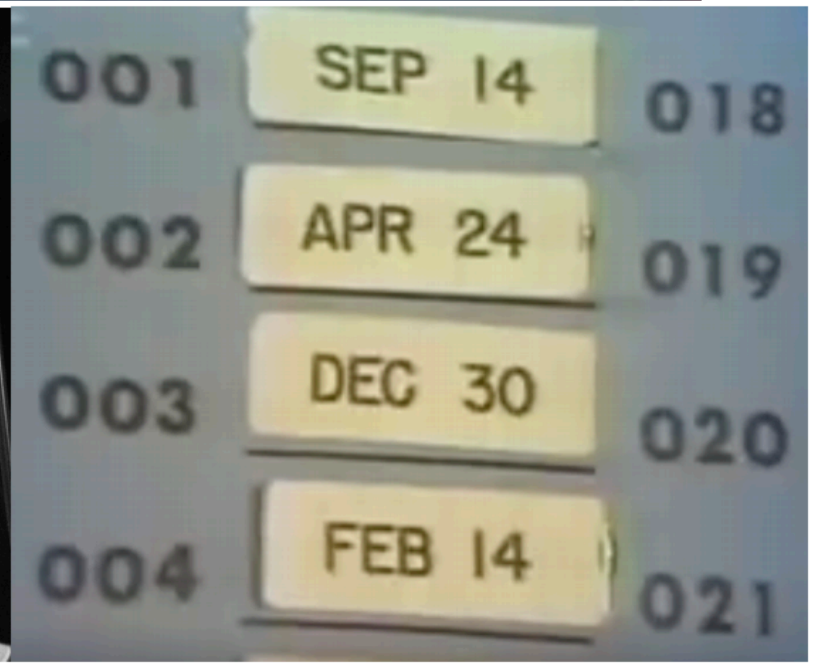
Using Quasi-Experimental Variation

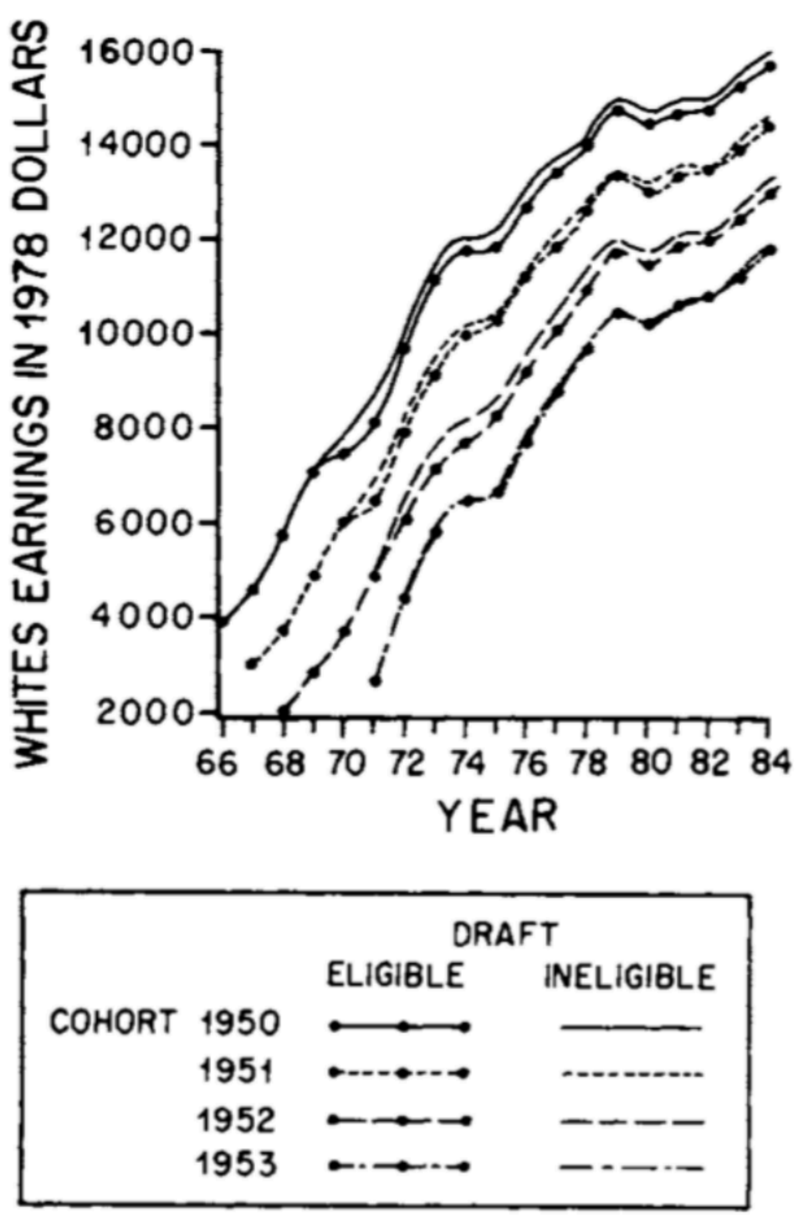
Arkansas	1996	1998	Difference
Benefit guarantee	$5,000	$4,000	–$1,000
Hours of work per year	1,000	1,200	200
Louisiana	**1996**	**1998**	**Difference**
Benefit guarantee	$5,000	$5,000	$0
Hours of work per year	1,050	1,100	50

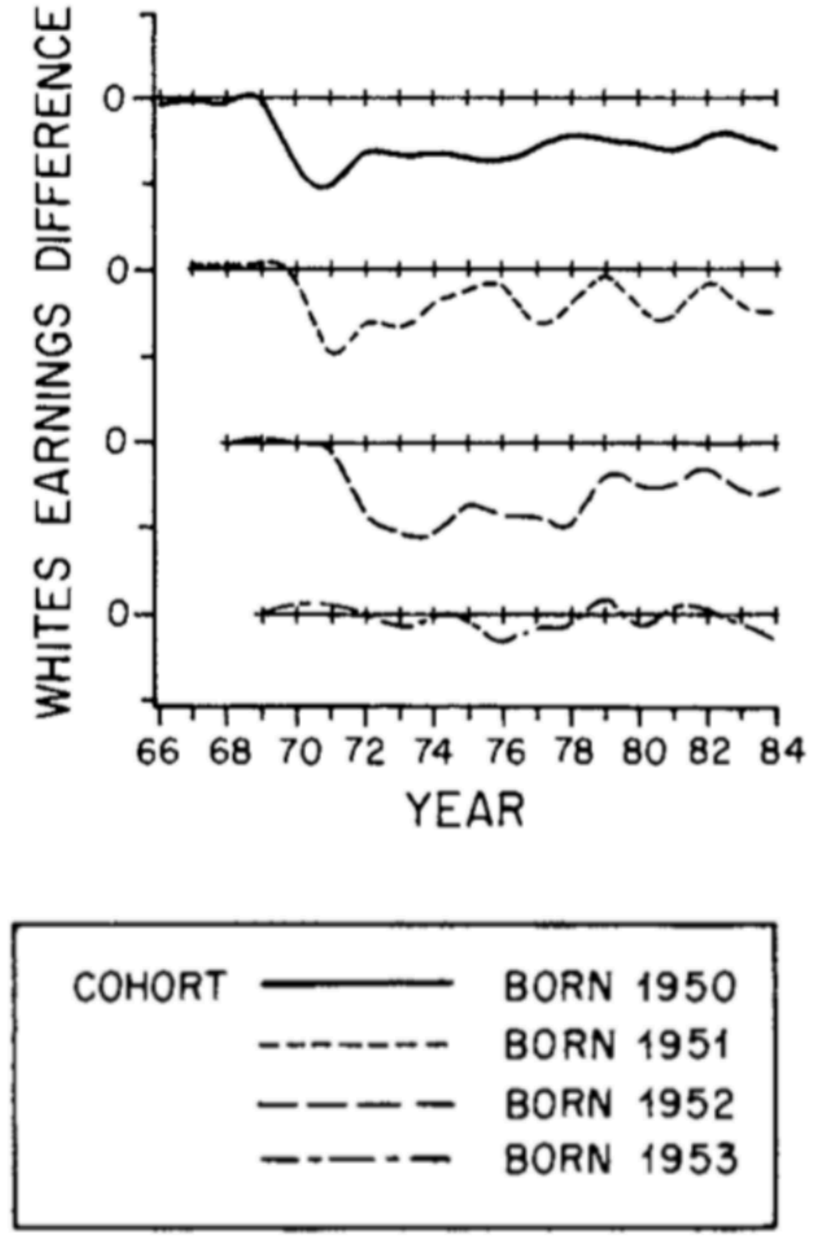
In Arkansas, there is a cut in the TANF guarantee between 1996 and 1998 and a corresponding rise in labor supply, so if everything is the same for single mothers in both years, this is a causal effect. If everything is not the same, we can perhaps use the experience of a neighboring state that did not decrease its benefits, Louisiana, to capture any bias to the estimates.

Quasi-Experiments
- Does rise in NJ minimum wage negatively affect employment?
- Draft to Vietnam war by lottery









If you get this, you are ready for exam