Public Finance II.

Lecture VIII - **Empirical Public Finance**

Matej Lorko
matej.lorko@euba.sk
www.lorko.sk

Office Hours (Room 5C.30)
Tue 10:00 – 10:45
Thu 12:30 – 13:15

Readings:

- Gruber, J. (2005). Public finance and public policy. Macmillan.
- Congdon, W. J., Kling, J. R., & Mullainathan, S. (2011). Policy and choice: Public finance through the lens of behavioral economics. Brookings Institution Press.

Why?

- Interestingly, it's a fact that highly intelligent women tend to marry men less intelligent than they are. Why do you think this might be?

Empirical public finance

- The central issue for any policy question is establishing a causal relationship between the policy in question and the outcome of interest.
  - Do lower welfare benefits cause higher labor supply among single mothers?
  - Does more pollution in the air cause worse health outcomes?
  - Do larger benefits for unemployment insurance cause individuals to stay unemployed longer?
- In this lecture, we discuss several approaches to distinguish causality from correlation. The gold standard for doing so is the randomized trial, which removes bias through randomly assigning treatment and control groups. Unfortunately, however, such trials are not available for every question we wish to address in empirical public finance. As a result, we turn to alternative methods such as time series analysis, cross-sectional regression analysis, and quasi-experimental analysis. Each of these alternatives has weaknesses, but careful consideration of the problem at hand can often lead to a sensible solution to the bias problem that plagues empirical analysis.

Empirical public finance

- **empirical public finance**: The use of data and statistical methods to measure the impact of government policy on individuals and markets.

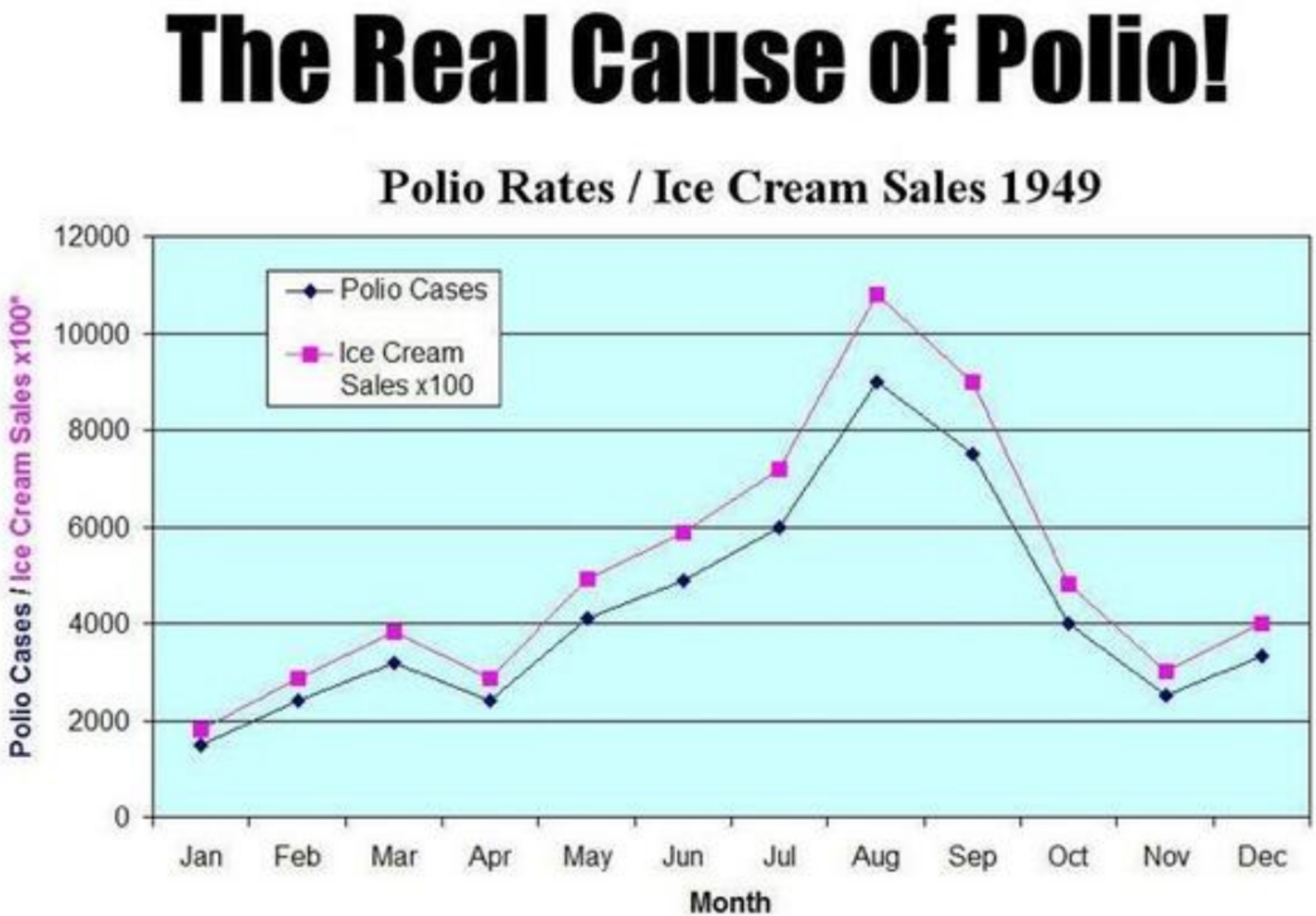
- fundamental issue faced by those doing empirical work in economics: disentangling causality from correlation.

- We say that two economic variables are **correlated** if they move together.

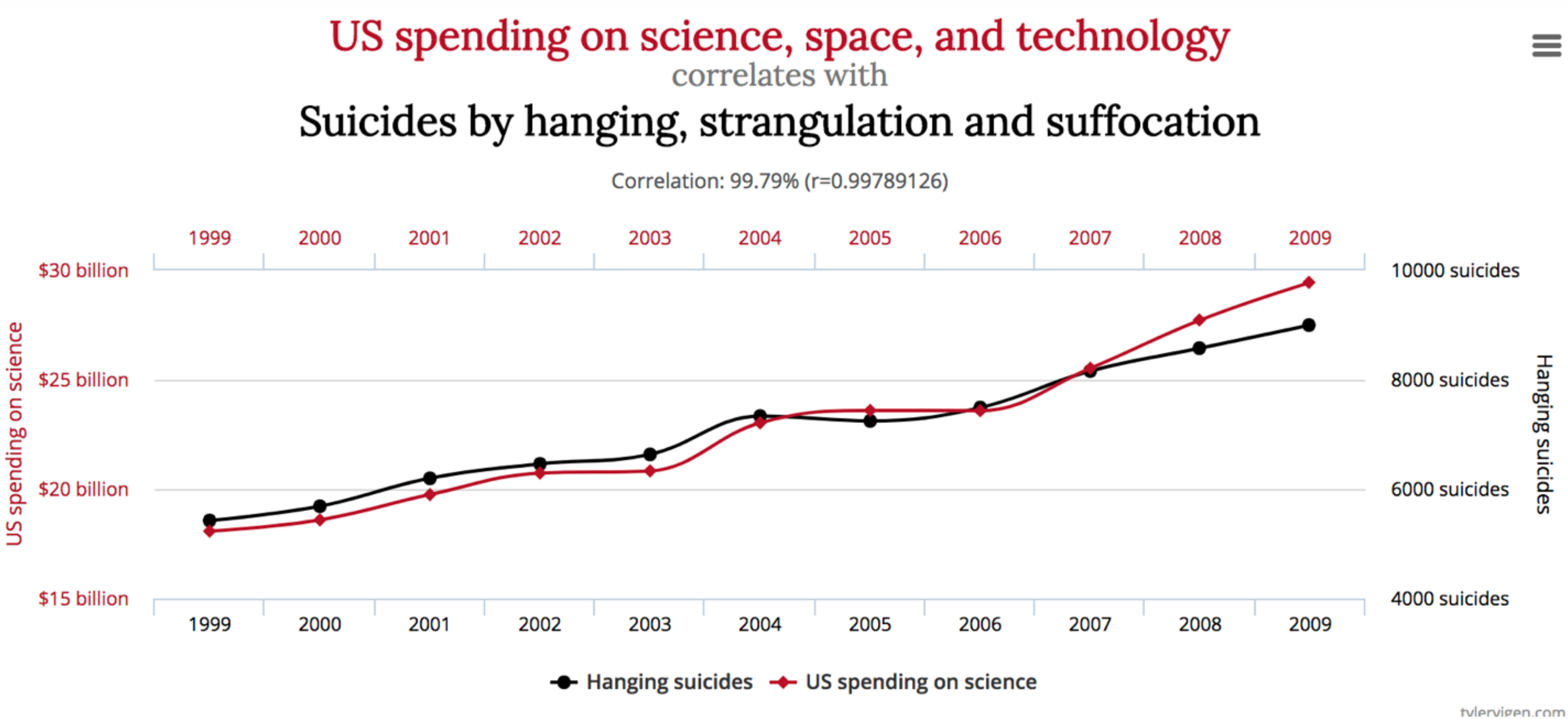
- But this relationship is **causal** only if one of the variables is causing the movement in the other.

- **identification problem**: given that two series are correlated, how do you identify whether one series is causing another?

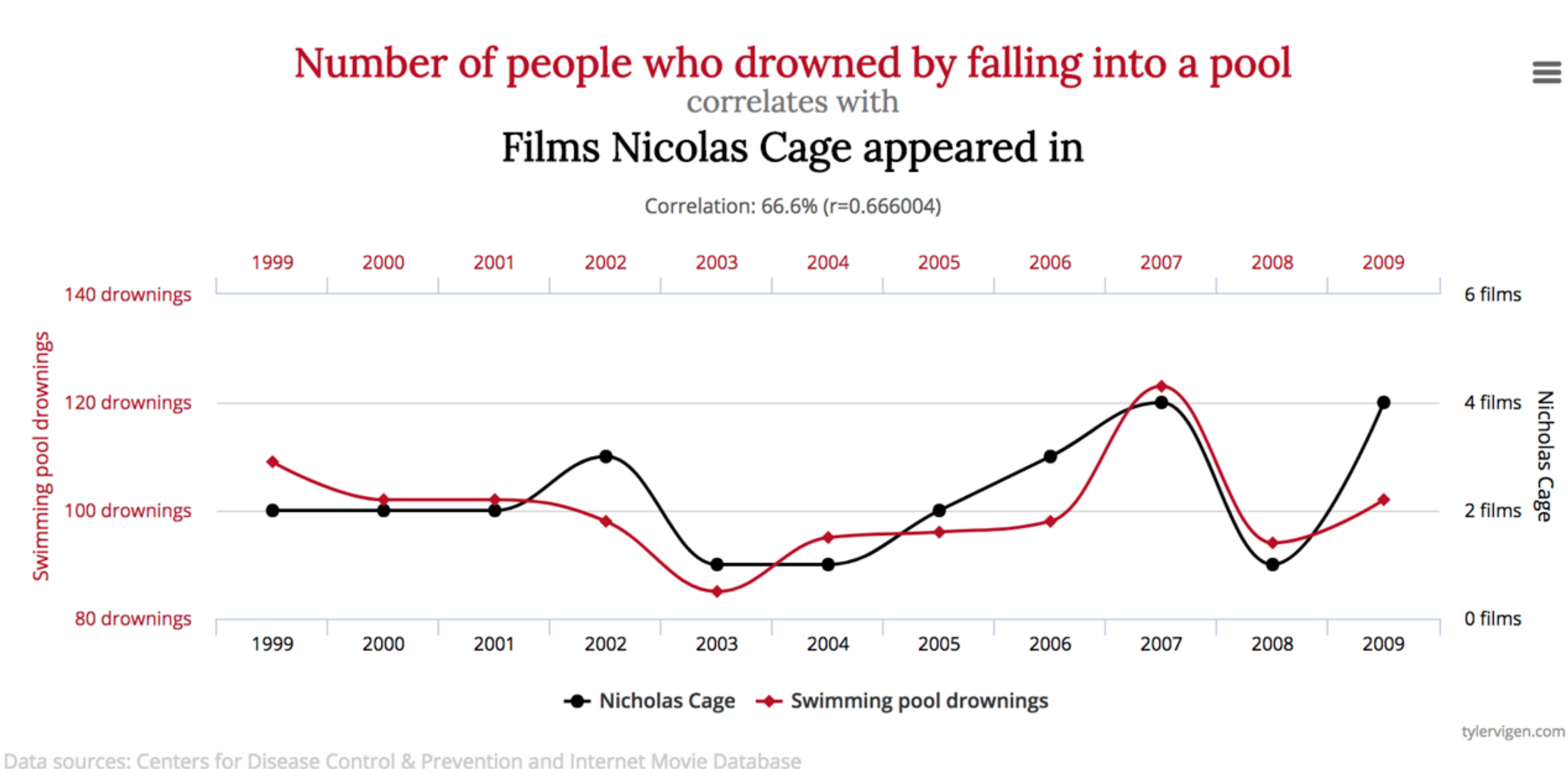
Correlation vs. Causation



Correlation vs. Causation



Correlation vs. Causation



Correlation vs. Causation
More examples

- Cholera in Russia
- SAT preparation courses vs. test scores
- Breast-feeding vs. malnutrition
- ERT
- U20 ice hockey rosters

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.

Assessing causatiom

- 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.

Assessing causatiom

- 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.

- This is a problem because for policy purposes what matters is causation. Policy makers typically want to use the results of empirical studies as a basis for predicting how government interventions will affect behaviors.

- Knowing that two factors are correlated provides no predictive power; prediction requires understanding the causal links between the factors. For example, the government shouldn't make policy based on the fact that breast-feeding infants are less healthy. Rather, it should assess the true causal effect of breast-feeding on infant health, and use that as a basis for making government policy.

- How can one draw causal conclusions about the relationships between correlated variables?

- A group of depressive kids was drinking one liter of coca cola daily. After a couple of months, their mental health got, on average better. Should we start treating depression with coca cola?

Golden standard for assessing causation: Randomized trials

- How can researchers address the problem of assessing causation? The best solution is through the gold standard of testing for causality: randomized trials.

- randomized trial: The ideal type of experiment designed to test causality, whereby a group of individuals is randomly divided into a treatment group, which receives the treatment of interest, and a control group, which does not.

- Randomized trials involve taking a group of volunteers and randomly assigning them to either a **treatment** group, which gets the treatment, or a **control** group, which does not. Effectively, **volunteers are assigned** to treatment or control **by the flip of a coin**.

- treatment group: The set of individuals who are subject to an intervention being studied.

- control group: The set of individuals comparable to the treatment group who are not subject to the intervention being studied.

Golden standard for assessing causation: Randomized trials

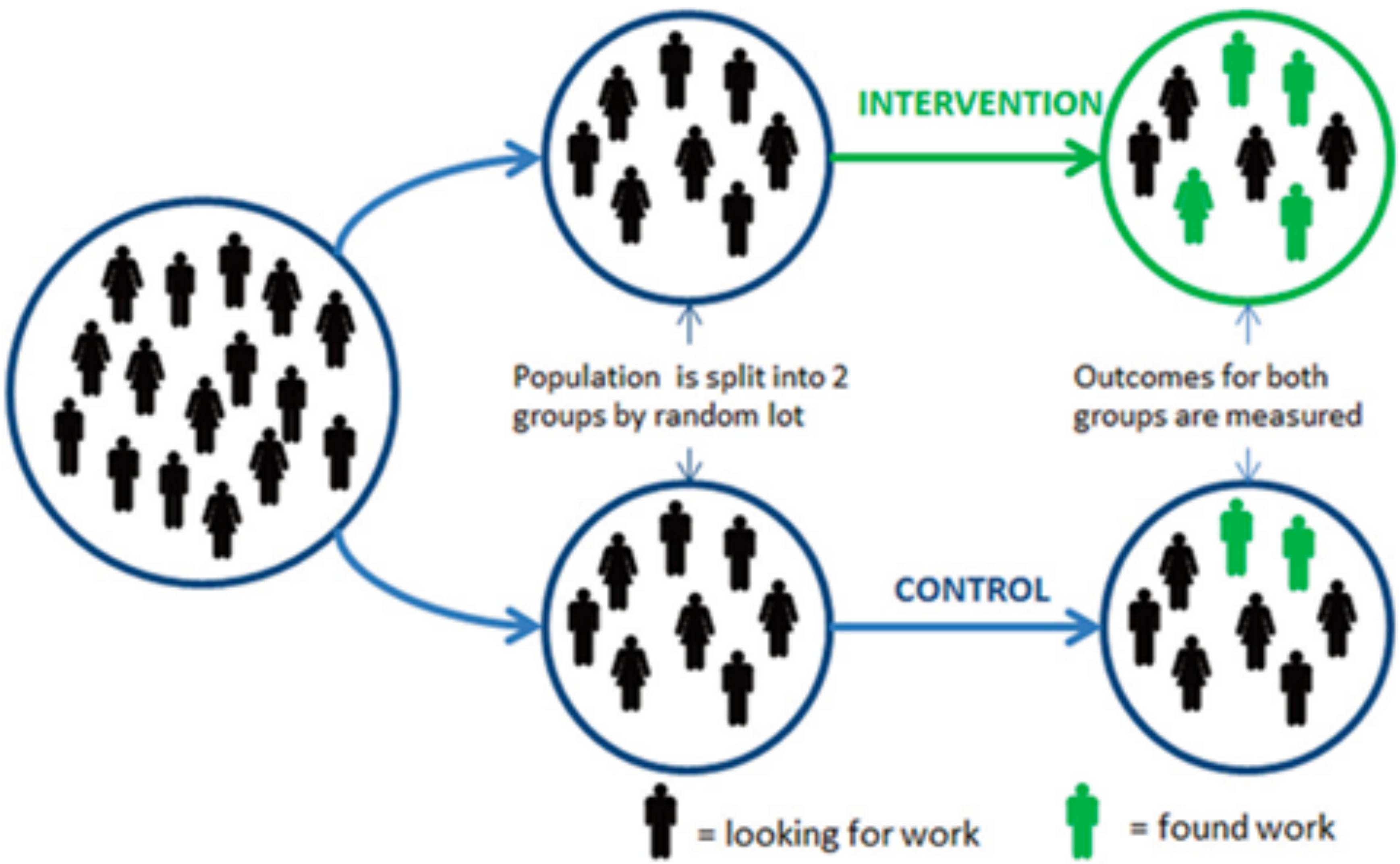


Figure 1. The basic design of a randomised controlled trial (RCT), illustrated with a test of a new 'back to work' programme.

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.

Possible problems with randomized trials

- There are a lot of questions for which executing randomized trial is not possible…
  - Do veterans have lower wages due to serving in war?
  - Do education raise wages?
  - What is the slope of a demand curve?
  - How does minimum wage affect the employment?
  - Does class size affect learning?
  - How much does alcohol raise the risk of crash?
  - Does access to information improve the market effectiveness?

Possible problems with randomized trials

- 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.

Estimating causation with observational data

- data from randomized trials are not always available when important empirical questions need to be answered. Typically, what the analyst has instead are observational data.

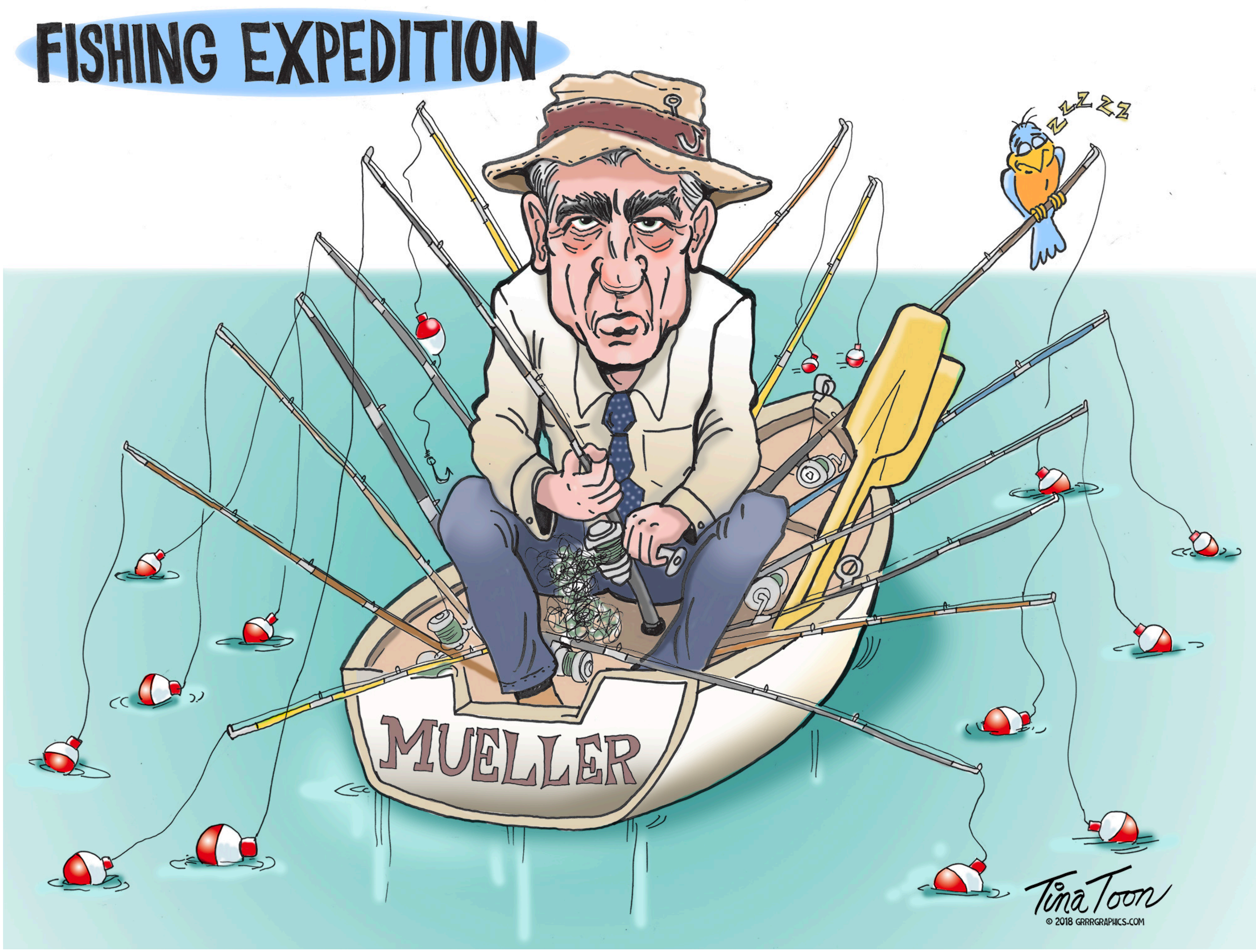
- observational data: data generated by individual behavior observed in the real world, not in the context of deliberately designed experiments.

- 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).

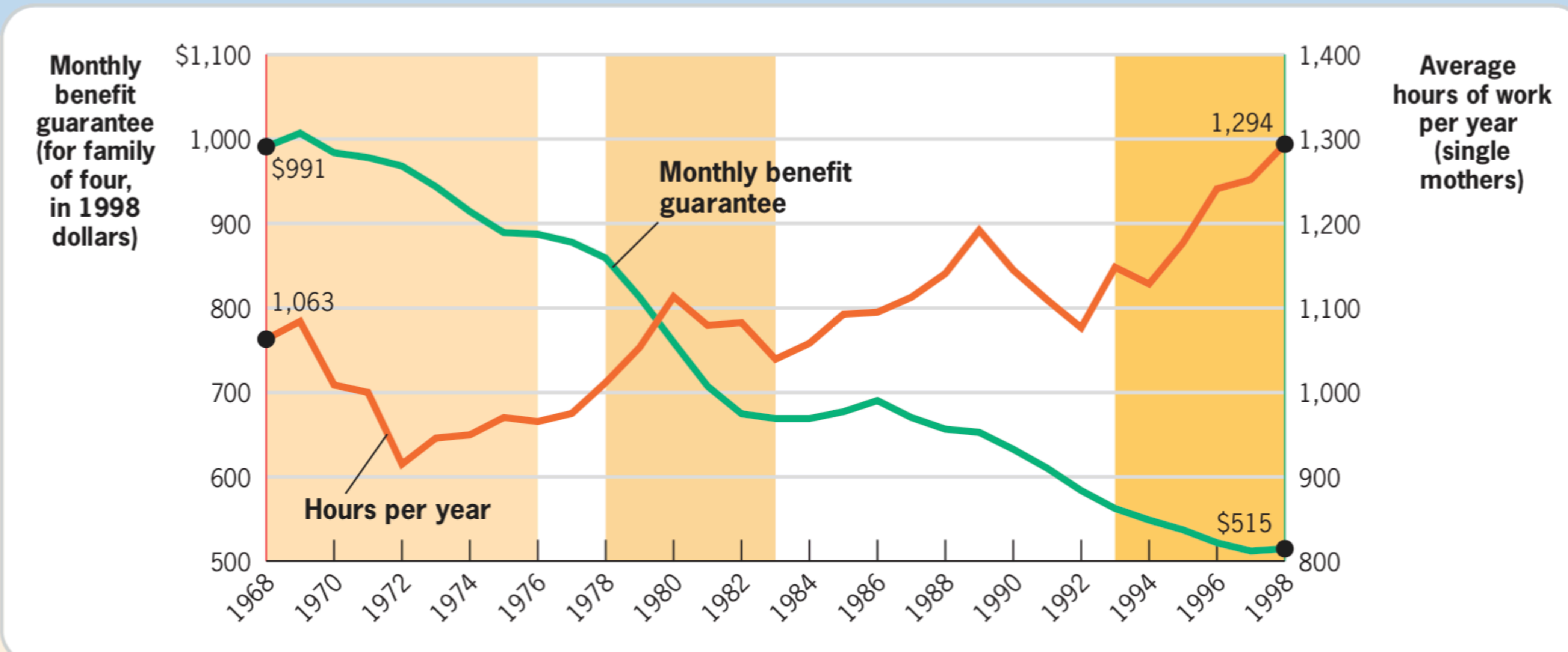
Bad hypothesis, bad result



Time Series Analysis

- One common approach to measuring causal effects with observational data is **time series analysis**, the analysis of the comovement of two series over time.

■ FIGURE 3-1



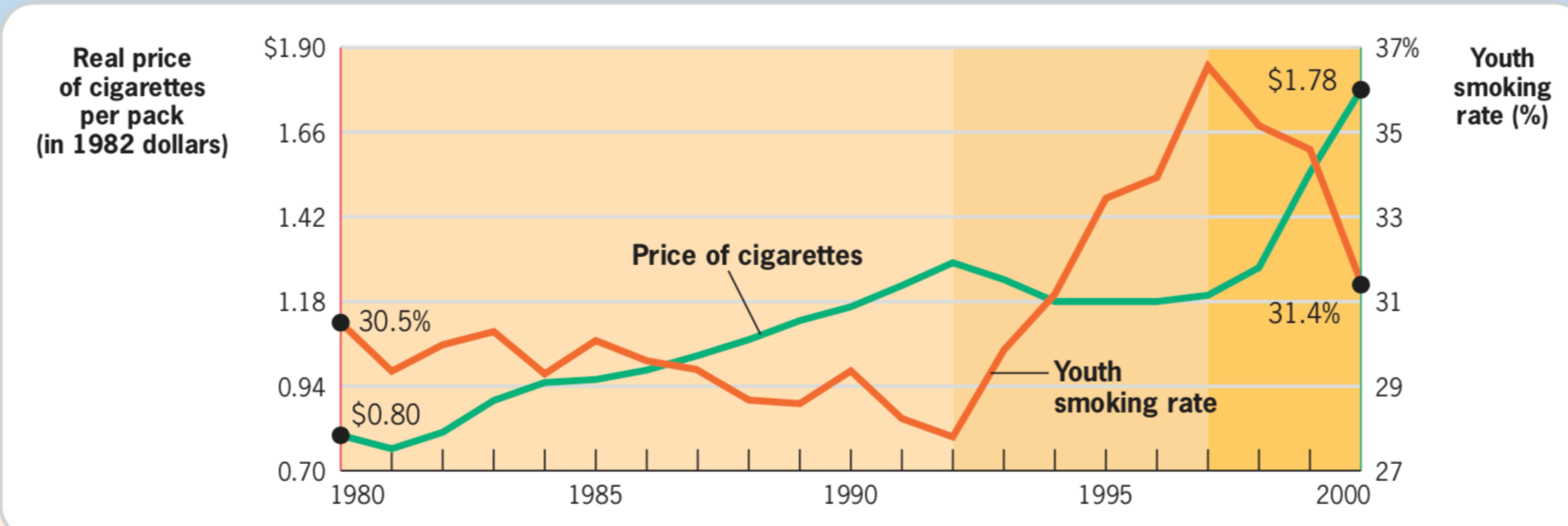
Average Benefit Guarantee and Single Mother Labor Supply, 1968–1998 • The left-hand vertical axis shows the monthly benefit guarantee under cash welfare, which falls from $991 in 1968 to $515 in 1998. The right-hand vertical axis shows average hours of work per year for single mothers, which rises from 1,063 in 1968 to 1,294 in 1998. Over this entire 30-year period, there is a strong negative correlation between the average benefit guarantee and the level of labor supply of single mothers, but there is not a very strong relationship within subperiods of this overall time span.

Time Series Analysis

- Problems with Time Series Analysis
    - Although the time series correlation can be striking, it does not necessarily demonstrate a causal effect. Other factors get in the way of a causal interpretation of the correlation over time; and these factors can cause bias in this time series analysis because they are also correlated with the outcome of interest.
- When Is Time Series Analysis Useful?
    - Is all time series analysis useless? Not necessarily. In some cases, there may be sharp breaks in the time series that are not related to third factors that can cause bias.
    - Thus, while time series correlations are not very useful when there are long-moving trends in the data, they are more useful when there are sharp breaks in trends over a relatively narrow period of time.

Time Series Analysis

■ FIGURE 3-2



Real Cigarette Prices and Youth Smoking, 1980–2000 • The left-hand vertical axis shows the real price of cigarettes per pack, which rises from $0.80 in 1980 to $1.78 in 2000. The right-hand vertical axis shows the youth smoking rate (the share of high school seniors who smoke at least once a month), which fell from 1980 to 1992, rose sharply to 1997, and then fell again in 2000 to roughly its 1980 level. There is a striking negative correspondence between price and youth smoking within subperiods of this era.

Source: Calculations based on data on smoking from Monitoring the Future survey and on tobacco prices from the Tobacco Institute.

Cross-Sectional Regression Analysis

- A second approach to identifying causal effects is **cross-sectional regression analysis**, a statistical method for assessing the relationship between two variables while holding other factors constant. By cross-sectional, we mean comparing many individuals at one point in time, rather than comparing outcomes over time as in a time series analysis.

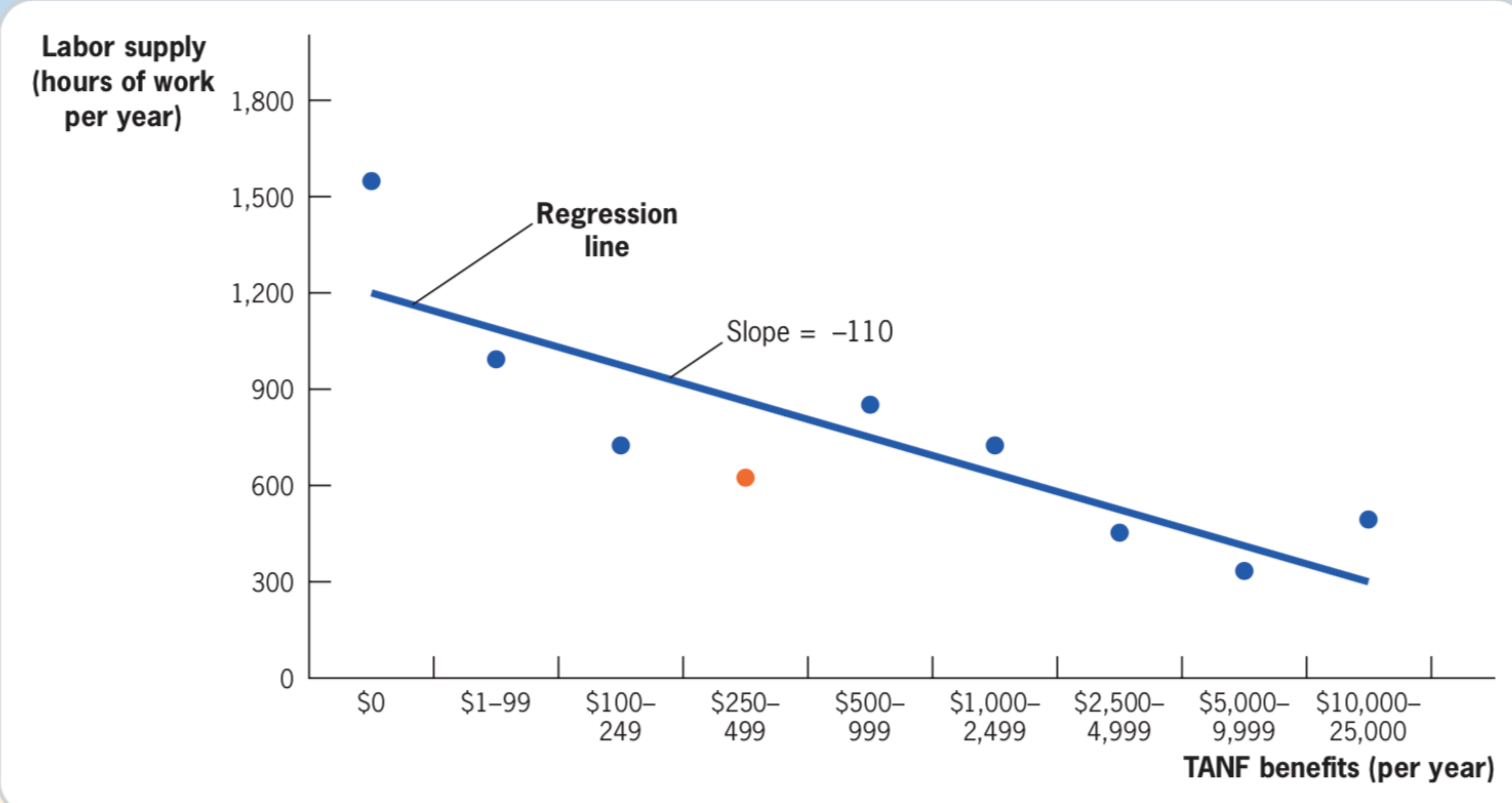
- Regression analysis describes (and quantifies) the relationship between the variable that you would like to explain (the dependent variable) and the set of variables that you think might do the explaining (the independent variables).

- The best approximation of such relationship is shown by the regression line. There is no single line that fits perfectly through this set of data points; instead, the linear regression finds the line that comes closest to fitting through the cluster of data points.

- Technically, this line is the one that minimizes the sum of squared distances of each point from the line. As a result, one major concern with linear regression analysis is outliers. An outlier, which is a point that is very far from the others, exerts a strong influence on this line, since we are minimizing the sum of squared distances, so a large distance has an exponentially large effect. For this reason, analysts often use other approaches that are less sensitive to such outlying observations.

Cross-Sectional Regression Analysis

■ FIGURE 3-4



TANF Benefit Income and Labor Supply of Single Mothers, Using CPS Data • Using data from the CPS, we group single mothers by the amount of TANF income they have. Those who are receiving the lowest level of TANF income are the ones providing the highest number of work hours.

Source: Calculations based on data from Current Population Survey's annual March supplements.

Cross-Sectional Regression Analysis

- The relationship between two variables approximated by the regression line is, again, not neccessarily causal. Therefore, we don't interrpret the results as "a x% reduction/rise in variable A is causing y% reduction/rise in variable B" but rather "a x% reduction/rise in variable A is associated with y% reduction/rise in variable B"

- Regression analysis has one potential advantage over correlation analysis in dealing with the problem of bias: the ability to include control variables. Control variables in regression analysis take into account other differences across individuals in a sample, so that any remaining correlation between the dependent variable and independent variable can be interpreted as a causal effect.

- However, in reality, control variables are unlikely to ever solve the problem of bias completely, as the key variables we want, are often impossible to measure in data sets. Usually, we have to approximate the variables we really want with what is available. These are imperfect proxies, however, so they don't fully allow us to control for differences.

Cross-Sectional Regression Analysis

- $y = \alpha + \beta x + e$
- where
  - α = constant (value for $x = 0$)
  - β = slope coefficient, represents the change y per unit change of x
  - e = error term, which represents the difference for each observation between its actual value and its predicted value based on the model

Reading regression analysis results

```
Linear regression                               Number of obs     =        130
                                                F(12, 117)        =       3.49
                                                Prob > F          =     0.0002
                                                R-squared         =     0.2353
                                                Root MSE          =      586.8
```

abserror	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
After	-252.8127	118.7671	-2.13	0.035	-488.0247	-17.60077
Before	-413.7776	111.5208	-3.71	0.000	-634.6386	-192.9165
Detail	283.7623	180.3411	1.57	0.118	-73.3938	640.9185
dem1	-1.658151	19.99805	-0.08	0.934	-41.26323	37.94693
dem2	15.81251	111.2361	0.14	0.887	-204.4847	236.1097
dem3	69.00018	50.26583	1.37	0.172	-30.54867	168.549
dem4	49.68547	53.67132	0.93	0.356	-56.60777	155.9787
dem6	-68.4688	52.44197	-1.31	0.194	-172.3274	35.38977
riskONE	-5.618447	25.47368	-0.22	0.826	-56.06774	44.83084
timeEST	-1.795241	1.531554	-1.17	0.244	-4.828403	1.237921
timeCONF	7.61295	10.98917	0.69	0.490	-14.15053	29.37643
confidence	-26.79666	74.69199	-0.36	0.720	-174.7202	121.1269
_cons	661.6392	568.5706	1.16	0.247	-464.3849	1787.663

Table 6: Linear regression analysis

Dependent variable	(1) Estimate	(2) Actual Duration	(3) Estimation bias	(4) Absolute estimation error
1. Info-After Treatment	323.63**	-94.88	380.79**	-252.81**
	(-140.63)	(-131.26)	(-161.68)	(-118.77)
2. Info-Before Treatment	216.79*	-124.96	316.48**	-413.78***
	(-127.3)	(-124.76)	(-150.28)	(-111.52)
3. Detailed Description Treatment	586.68**	-8.68	526.98*	283.76
	(-266)	(-146.65)	(-267.93)	(-180.34)
4. Age	-5.54	5.46	-10.35	-1.66
	(-23.4)	(-19.59)	(-25.66)	(-20)
5. Female	-84.79	-88.01	13.1	15.81
	(-153.07)	(-96.09)	(-174.69)	(-111.24)
6. Self-reported math skill	27.8	-188.03***	212.58***	-69
	(-79.59)	(-53.23)	(-79.99)	(-50.27)
7. Current degree of study	34.51	1.51	28.97	49.69
	(-79.7)	(-41.54)	(-72.4)	(-53.67)
8. Employment status	52.91	-10.26	57.01	-68.47
	(-68.78)	(-47.06)	(-78.08)	(-52.44)
9. Risk attitudes	45.22	28.89	11.06	-5.62
	(-30.23)	(-26.61)	(-35.89)	(-25.47)
10. Time spent estimating	-2.83	-1.73	-0.77	-1.8
	(-2.37)	(-1.26)	(-2.65)	(-1.53)
11. Time spent indicating confidence in estimate	10.72	11.24	-1.77	7.61
	(-15.37)	(-12.69)	(-19.28)	(-10.99)
12. Subjective confidence in estimate	-243.00**	45.3	-259.98**	-26.8
	(-97.14)	(-77.81)	(-119.6)	(-74.69)
13. Estimate		0.12**		
		(-0.06)		
Constant	1317.07**	206.68	956.86	661.64
	(-569.07)	(-646.97)	(-853.56)	(-568.57)
N	130	130	130	130
R^2	0.13	0.20	0.12	0.24

Note: Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1%-level, respectively.

Quasi-Experiments

- As noted earlier, public finance researchers cannot set up randomized trials and run experiments for every important behavior that matters for public policy. We have examined alternatives to randomized trials such as time series and cross-sectional regression analysis, but have also seen that these research methods have many shortcomings which make it hard for them to eliminate the bias problem.

- Is there any way to accurately assess causal influences without using a randomized trial? Is there an alternative to the use of control variables for purging empirical models of bias?

- Over the past two decades, empirical research in public finance has become increasingly focused on one potential middle-ground solution: the quasi-experiment, 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

Quasi-Experiments

TABLE 3-1

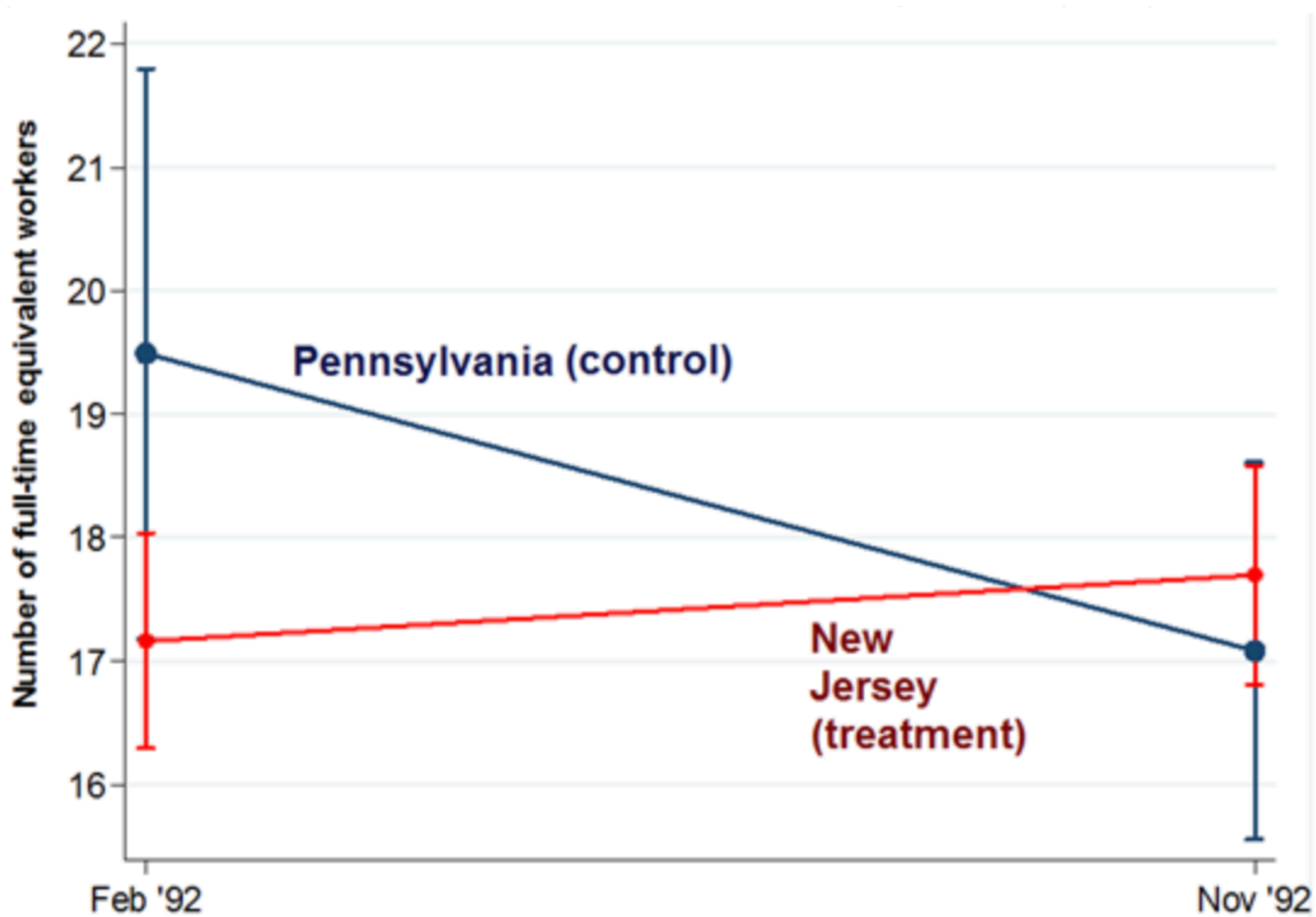
Using Quasi-Experimental Variation

Arkansas	1996	1998	Difference
Benefit guarantee	\$5,000	\$4,000	–\$1,000
Hours of work per year	1,000	1,200	200

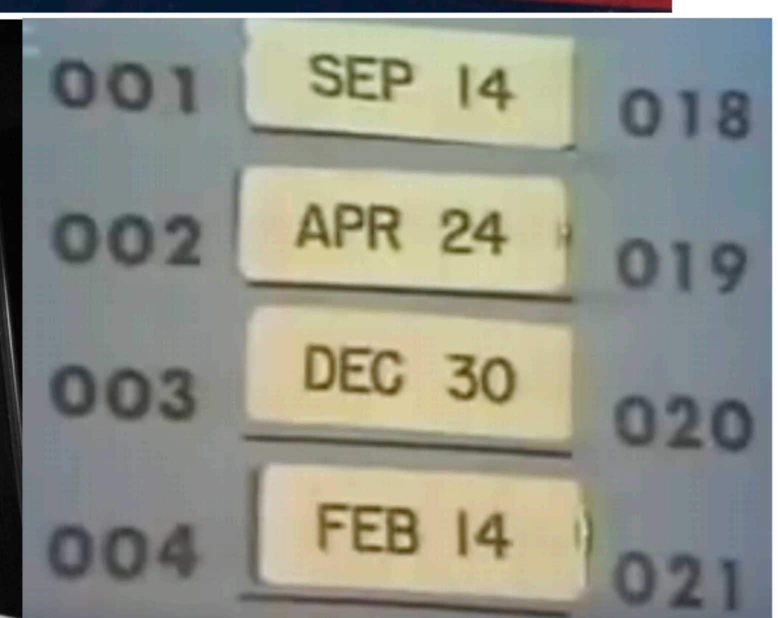
Louisiana	1996	1998	Difference
Benefit guarantee	\$5,000	\$5,000	\$0
Hours of work per year	1,050	1,100	50

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.

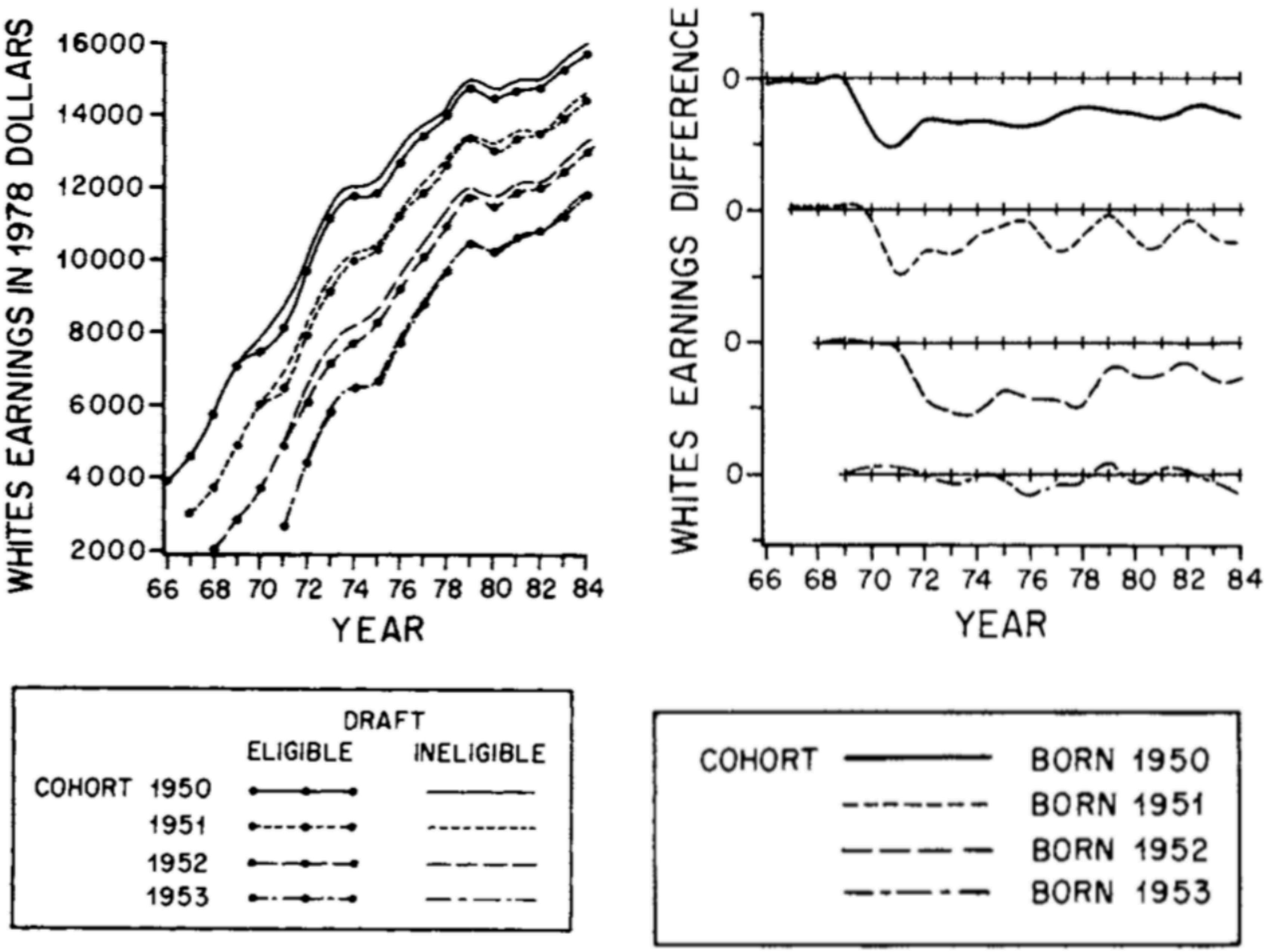
Quasi-Experiments - does rise in NJ minimum wage negatively affect employment?



Quasi-Experiments - Draft to Vietnam war by lottery



Quasi-Experiments - Draft to Vietnam war by lottery



Structural Modeling

- The randomized trials and quasi-experimental approaches previously described have the distinct advantage that, if applied appropriately, they can address the difficult problem of distinguishing causality from correlation.Yet they also have two important limitations. First, they only provide an estimate of the causal impact of a particular treatment. Say that an experiment found that cutting variable A by 15% raised variable B rates by 4.5 percentage points. This is the best estimate of the impact of cutting variable A by 15%, but it may not tell us much about the impact of cutting variable A by 30%, or of raising variable A by 15%. That is, we can't necessarily extrapolate from a particular change in the environment to model all possible changes in the environment. These approaches give us a precise answer to a specific question, but don't necessarily provide a general conclusion about how different changes in something might affect behavior.

- The second limitation is that these approaches can tell us how outcomes change when there is an intervention, but often they cannot tell us why. We often care about the **structural estimates** of responses, the estimates that tell us about features of utility that drive individual decisions, such as substitution and income effects. Randomized or quasi-experimental estimates provide **reduced form estimates** only.

- Reduced form estimates show the impact of one particular change on overall responses. This second disadvantage of randomized or quasi-experiments is thus related to the first: if we understood the underlying structure of responses, it might be possible to say more about how people would respond to different types of policy interventions.

Structural Modeling

- These issues have led to the vibrant field of structural estimation. Using this research approach, empirical economists attempt to estimate not just reduced form responses to the environment but the actual underlying features of utility functions. They do so by more closely employing the theory to develop an empirical framework that not only estimates overall responses, but also decomposes these responses into, for example, substitution and income effects.

- Structural models potentially provide a very useful complement to experimental or quasi-experimental analyses. Yet structural models are often more difficult to estimate than reduced form models because both use the same amount of information, yet structural models are used to try to learn much more from that information.

Quantitative support for causation

- Model does not have to be truthful, but it has to be useful
- The more of the following conditions are met, the easier is to assign causation
  - Effect is significant - not only statistically significant, but also practically significant
  - Effect is consistent - same or similar results were achieved by other studies
  - Effect is specific - the factor of intereest influences only the response of interest, not also 10 other things
  - Effect follows time continuity - if x causes y, then x has to happen first
  - Effect is monotonous - more x causes more change in y
  - Effect is plausible - we know the mechanism behind the effect
  - Effect if supported by experimental results

Takeaways

- A primary goal of empirical work is to document the causal effects of one economic factor on another.
- The difficulty with this goal is that it requires treatment groups (those who are affected by policy) and control groups (those not affected) who are identical except for the policy intervention.
  If these groups are not identical, there can be bias— that is, other consistent differences across treatment/control groups that are correlated with, but not due to, the treatment itself.
- Randomized trials are the gold standard to surmount this problem. Since treatments and controls are identical by definition, there is no bias, and any differences across the groups are a causal effect.
- Time series analysis is unlikely to provide a convincing estimate of causal effects because so many other factors change through time.
- Cross-sectional regression analysis also suffers from bias problems because similar people make different choices for reasons that can't be observed, leading once again to bias. Including control variables offers the potential to address this bias.
- Quasi-experimental methods have the potential to approximate randomized trials, but control groups must be selected carefully in order to avoid biased comparisons.

- Your state introduced a tax cut in the year 1999. You are interested in seeing whether this tax cut has led to increases in personal consumption within the state. You observe the following information:

Year	Consumption in your state
1994	300
1996	310
1998	320
2000	350

- a. Your friend argues that the best estimate of the effect of the tax cut is an increase in consumption of 30 units, but you think that the true effect is smaller, because consumption was trending upward prior to the tax cut. What do you think is a better estimate?

- b. Suppose that you find information on a neighboring state that did not change its tax policy during this time period.You observe the following information in that state:

Year	Consumption in neighboring state
1994	260
1996	270
1998	280
2000	300

- Given this information, what is your best estimate of the effect of the tax cut on consumption? What assumptions are required for that to be the right estimate of the effect of the tax cut? Explain.

If you get this, you are ready for exam