research design

adam okulicz-kozaryn

adam.okulicz.kozaryn@gmail.com

this version: Wednesday 31st January, 2018 18:06

outline

intuition

research design basics

DID and program evaluation (Wheelan, 2013, ch13)

level of analysis

outline

statistics

- "statistics is the science of learning from data"
- "the science of collecting and analyzing data for the purpose of drawing conclusions and making decisions."
- good data are the key! GIGO (Wheelan, 2013, ch7)
- study what you are interested in, and what you are knowledgable about!

example: use data to disprove your convictions!

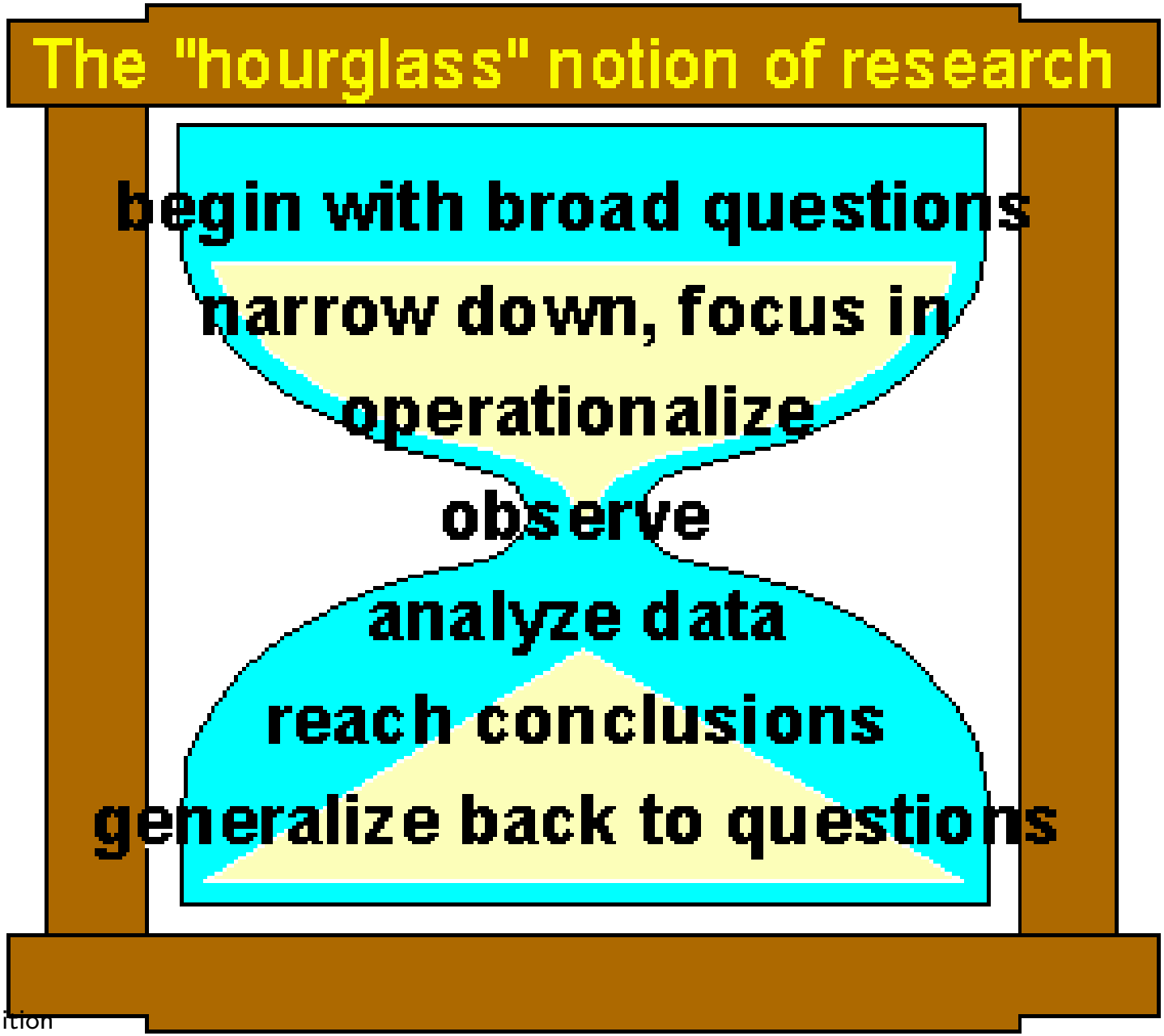
- i knew it by heart that cities are places of largest inequalities
- `https://projects.newyorker.com/story/subway`

- but so are unequal rural areas!
- `https://www.google.com/search?q=pew+inequalit+by+county&ie=utf-8&oe=utf-8`
- correlation of pop siz and gini just .1

setup

- ◇ either critique research or do it yourself:
  - design the problem (a question, research idea, hypothesis)
  - collect the data (download)
  - summarize the data (descriptive statistics)
  - analyze (inferential statistics), interpret, communicate
- ◇ many people just analyze;
  - you need to interpret–what does it mean
- ◇ and you need to communicate
  - that is interpret in the most simple way possible
  - most people don't understand statistics

the hourglass (Trochim)



begin with broad questions

- ideally study something that you are passionate about
- (again, whatever that is, you will find good data about it)
- yet, ieally connect to existing literature
- know what has been found (goog scholar)
- and built on it

narrow down, focus in

- be specific about what exactly/what aspect
- you are looking at...
- too broad ideas cannot be tested
- it has to be as specific as possible
- if needed you may break it down into several specific hypotheses
- anyone having any hypotheses? give me few examples?

operationalize: have a hypothesis

- hypothesis: brief and clear statement that can be tested
  - measured with variables and specified "+" or "-" effect
- express your idea in observable/measurable terms
- translate words/idea into a mathematical relationship
- eg increase in X is associated with decrease in Y
  - where X and Y are specific variables
  - say, income increases happiness
- and then use research methods, interpret results
  - and answer initial questions

do not over complicate

- students have a tendency to over complicate...
  - to test grand theories; eg what causes economic growth
  - it takes years, generations to answer those grand questions
- start simple, less ambitious
  - you can always extend it if you have time/resources
  - it is better to finish a simple project than not to finish a complicated one
- and much easier and faster to contribute locally than academically

outline

intuition

research design basics

DID and program evaluation (Wheelan, 2013, ch13)

level of analysis

qualitative vs quantitative

- much of the following applies whether you do qualitative or quantitative research
- research design is a class itself
- we will cover only basics; for more:
- `http://www.socialresearchmethods.net/kb/design.php`

external validity (Wheelan, 2013, ch10)

- external validity is about generalizability
- can i say something about RU in general by analyzing you?
- how about just RU-Camden ?
- no ! people at Law school, computational biology are likely to be different
- and even per PA, I would ideally like to have a random sampling
- note: random sampling is different from randomization or random assignment (experiment)
- `http://knowledge.sagepub.com/view/researchdesign/n146.xml#n146`

internal validity

- internal validity is about causality
- you have internal validity if you can claim that X causes Y
  - eg some drug X causes some disease Y to disappear
  - `http://knowledge.sagepub.com/view/researchdesign/n43.xml#n43`
  - `http://knowledge.sagepub.com/view/researchdesign/n192.xml#n192`

causality

- much of research design is about causality
  - want to show X $\rightarrow$ Y
- correlation is necessary for causality
  - (rarely suppressor var makes it unnecessary, eg (Mazur, 2011))
  - but not sufficient (eg `http://www.tylervigen.com/`)
- careful! humans have illusion of causality:
  tend to see causality where there is none!
- `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4488611`
- `http://onlinelibrary.wiley.com/doi/10.1348/000712610X532210`

INUS condition (Mackie, 1980)

- very useful way of thinking about causality:
  Insufficient but Non-redundant part of
  Unnecessary but Sufficient Condition
- most causes are INUS conditions
- eg a cigarette as a cause of forrest fire
  - it's Insufficient, because by itself it is not enough, eg you also need oxygen, dry leaves, etc
  - it is contributing to fire, hence Non-redundant
- and **along with other stuff** (oxygen, dry leaves etc) it constitutes Unnecessary but Sufficient Condition
  - it's not necessary for fire, it can be lightening, etc
  - but it's sufficient – it's enough to start the fire

INUS condition

- IN is your X
- US is set of X's (your X+other X's)
- the bottom line is that there are always:
  - multiple alternative causes
  - and multiple steps in causal process
- or you could say there is a train of causality:
  - multiple things have to happen for outcome to occur
- say airplane fall: multiple things had to happen:
  - pilot, traffic control, weather, etc
- same with everything: career success, marriage, etc

basic concepts

- Y: a dependent variable, outcome
- X: an independent variable, predictor
- (T: (treatment), like X)
- Z: some other variable
- want to show $X \rightarrow Y$; X affects (causes) Y
- and not the other way round $(Y \rightarrow X)$
- and not $Z \rightarrow Y$; eg X(CO_2),Y(temp), Z(sun temp)
- it is difficult to argue !
- after all, there are unknown unknowns (Z's that we are unaware of)

The Problem: Unknown Unknowns (D Rumsfeld)

- there are known knowns; there are things we know that we know
- there are known unknowns; that is to say, there are things that we now know we don't know
- but there are also unknown unknowns
  –there are things we do not know we don't know
- how do we deal with unknown unknowns?
- do an experiment!
- bottom line: be humble and sceptical/suspicious
- eg: unknown unknown: value of (need for) specific discipline degree in the future

the gold standard [ask IRB appr!]

- the experimental design give few examples
- only with experiment can confidently argue causality
- and it is because randomization takes care of the known and unknown predictors of the outcome
  (draw a picture of 2 groups of people)
- in other words, it establishes a counterfactual (next slides)
- but wait !
- most of the time we cannot have an experimental design because it is unethical and politically impossible
  eg we cannot randomly assign kids to bad school or to smoking `http://www.socialresearchmethods.net/kb/desexper.php`

causality without experiment?

- yes! well maybe, but you need to do lots of work...
- essentially you want to exclude alternative explanations
- so you act like a devil's advocate...
- and try to abolish your story / find an alternative explanation
- if you cannot find any, then your story is right ...
  - until disproved
  - all swans are white...untill...you find a black one!

The Problem put another way: Counterfactual

- esentially need to compare:
  - what happened to the outcome (Y) due to the treatment (T or X)
  - to what would have been (Y), had the treatment not happened
- eg we got a new teacher and now kids perform better on SAT
  - to know whether the teacher caused better performance we would need to know what would have happened to SAT scores without this teacher (scores might have gone up due to Z),
  - and compare it to what actually happened

The Problem put another way: Counterfactual

- the problem is that we do not observe counterfactual (we can try to infer it though)
- counterfactual is the effect of all knowns/unknowns
  - (incl. unknown unknowns)
- how do we deal with lack of counterfactual
- do an experiment!
- (or if you cannot, try to estimate or infer it somehow)

threats to internal validity

- can still argue causality, but think about threats!
- time: history, maturation, regression to the mean
  - things develop over time in a certain way
- selection bias, self selection
  - does smoking causes cancer ?
  - maybe less healthy people select to smoke ?
- something else (Z) happened that caused Y
- reverse causality
- `http://knowledge.sagepub.com/view/researchdesign/n192.xml#n192`

reverse causality OR chicken-egg dilemma

- you may try to find some other X that measures the same or similar concept and that cannot be caused by Y
- eg instead of education $\rightarrow$ wage; do father's education$\rightarrow$ wage (your wage can reverse cause your education, but not your father's education)
- find some exogenous (external) shock: policing$\leftrightarrow$crime
- but terror attack/alert $\rightarrow$policing$\rightarrow$crime; we know that policing$\rightarrow$crime; not the other way round
  - `https://www.law.upenn.edu/fac/jklick/48JLE267.pdf`
- or dating happiness–which comes first? happy folks more likely to be dated!

natural experiment

- again most of the time you cannot have an experiment
- but there are natural experiments or exogenous shocks
- exogenous meaning that they are caused externally (like an experimenter's randomization) and somewhat randomly (at least with relation to a problem at hand
- eg earthquake (any weather, eg storm); terrorist attack; policy change (less random)

PRE, POST or BEFORE, AFTER

- one treatment group over time

  eg you can trace unemployment over time in Camden

  draw interrupted time series

- and, say, you can find that it increased during Reagan administration...
- still, you cannot argue causality right away !
- there may be lots of alternative explanations, eg shift away from manufacturing during the same time, etc etc

TREAT, CONTROL or EXPERIMENT, COMPARE

- compare across 2 groups over time
- eg you can compare crime in Camden and Newark, while there was some intervention in Camden only, say new policing approach
- draw 4 boxplots: Camden, Newark: before after
- see blood pressure in graphs:
- `http://www.ats.ucla.edu/Stat/stata/library/GraphExamples/code/grbox1.htm`
- (or see actual paper–scroll down to box plots in the middle `http://www.bmj.com/content/341/bmj.c3215`)

examples of designing research

- say a mayjor employer comes in,
- say Subaru in its block group
- or Salvation Army in its block group (Crammer Hill)
- look at housing prices (can proxy economic development)
- `https://www.zillow.com/research/data/`
- or see race by bl gr in the are
  `https://www.census.gov/2010census/popmap/`
- you can get so many variables at bl gr level!
- lets see social explorer and bl gr lev vars

>>>probably stop here and pick up next wk

- or before designs shown in graphs
- everyone is different! many ways to excel, eg opportunities for extra credit if you fail assignments: few will excel by doing math; few will excel by doing civic engagment; few will excel by having great research ideas–if you are motivated you will excel one way or the other!

ex post facto:

- very common...it is *no* design
- non-experimental, cross-sectional, observational, correlational; you'll most likey do or read this
- we start investigation "after the fact"
- no time involved, don't know whether X precedes Y
- both, X and Y are observed at the same time examples?
- (but X must precede Y in order to be causal)
- practically impossible to argue causality here
- but cheap and big N, and good external validity

ex post facto:

- useful, many "causes" were discovered using observational studies
- eg smoking→cancer was found out using ex post facto
- and then confirmed using better designs
- `http://knowledge.sagepub.com/view/researchdesign/n145.xml`
- `http://knowledge.sagepub.com/view/researchdesign/n271.xml#n271`

before-after (pre-post): blackboard: schematic

- measured Y, then do X, and then measured Y again
- eg measured readership at the library , buy some cool stats books ; measured readership again
- eg measured crime rate , put more police on the streets ; measured crime again
- eg measured soup consumption , changed soup ; measured soup consumption again
- anyone did pre/post? eg working at school?
- tried new programs, new approaches?
- or simply pre-post without T, say to identify highest and lowest gain students

(two group) comp chng: blackbrd: schematic

- eg Hypothesis: police with better guns fights crime better
- measured crime rate in 2010 in Camden and Newark
  - in 2011 give super guns to police in Camden , (but not in Newark)
  - in 2012 measured crime rate Camden and Newark
- if crime rate dropped more in Camden than in Newark, then we have evidence that the guns worked

interrupted time series:

- eg the new anti-unemployment program in Camden decreased unemployment
- get data about unemployment in Camden from 1990 to 2010
- say the unemployment program began in 2001
- produce a time series plot (mark a vertical line in 2001: intervention/treatment)
- if there was a change in trend after 2001, we conclude that the program worked

interrupted time series:

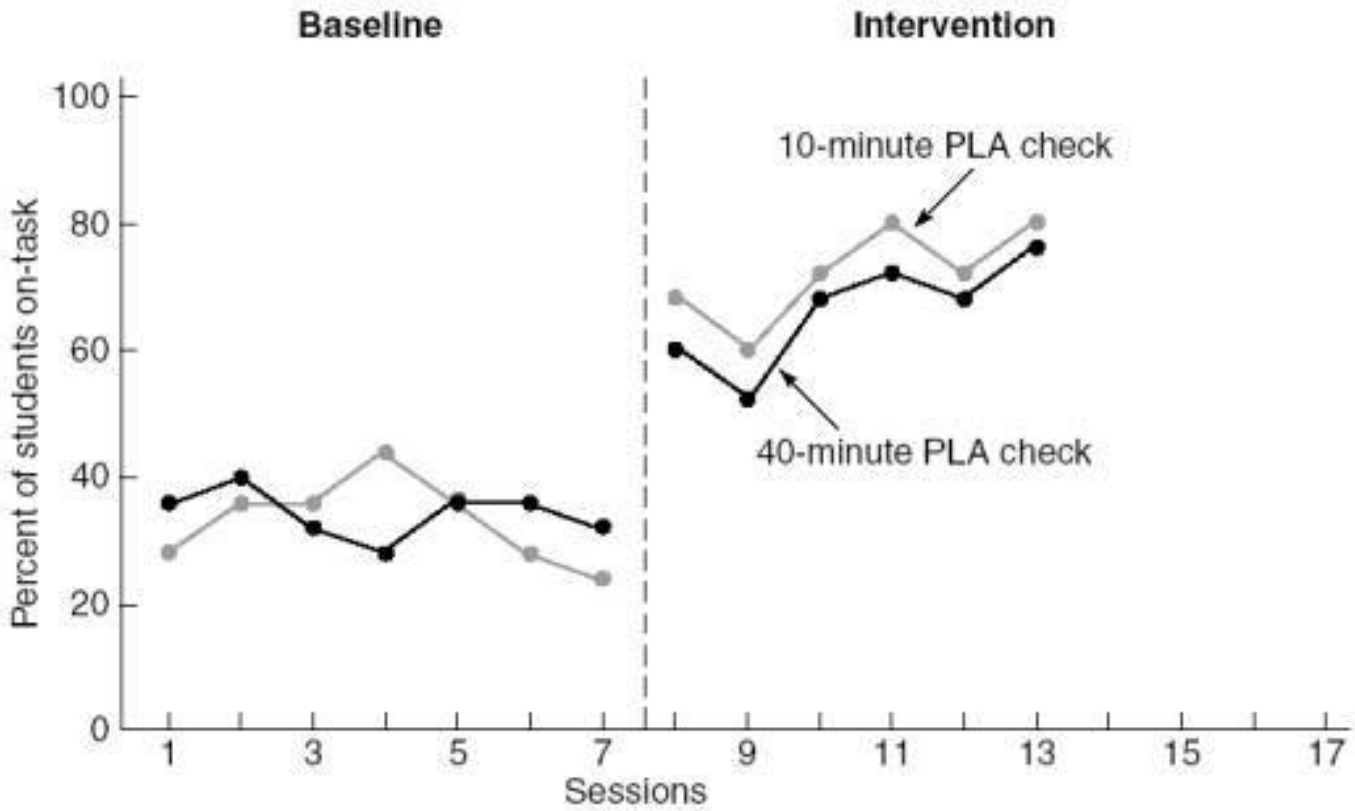
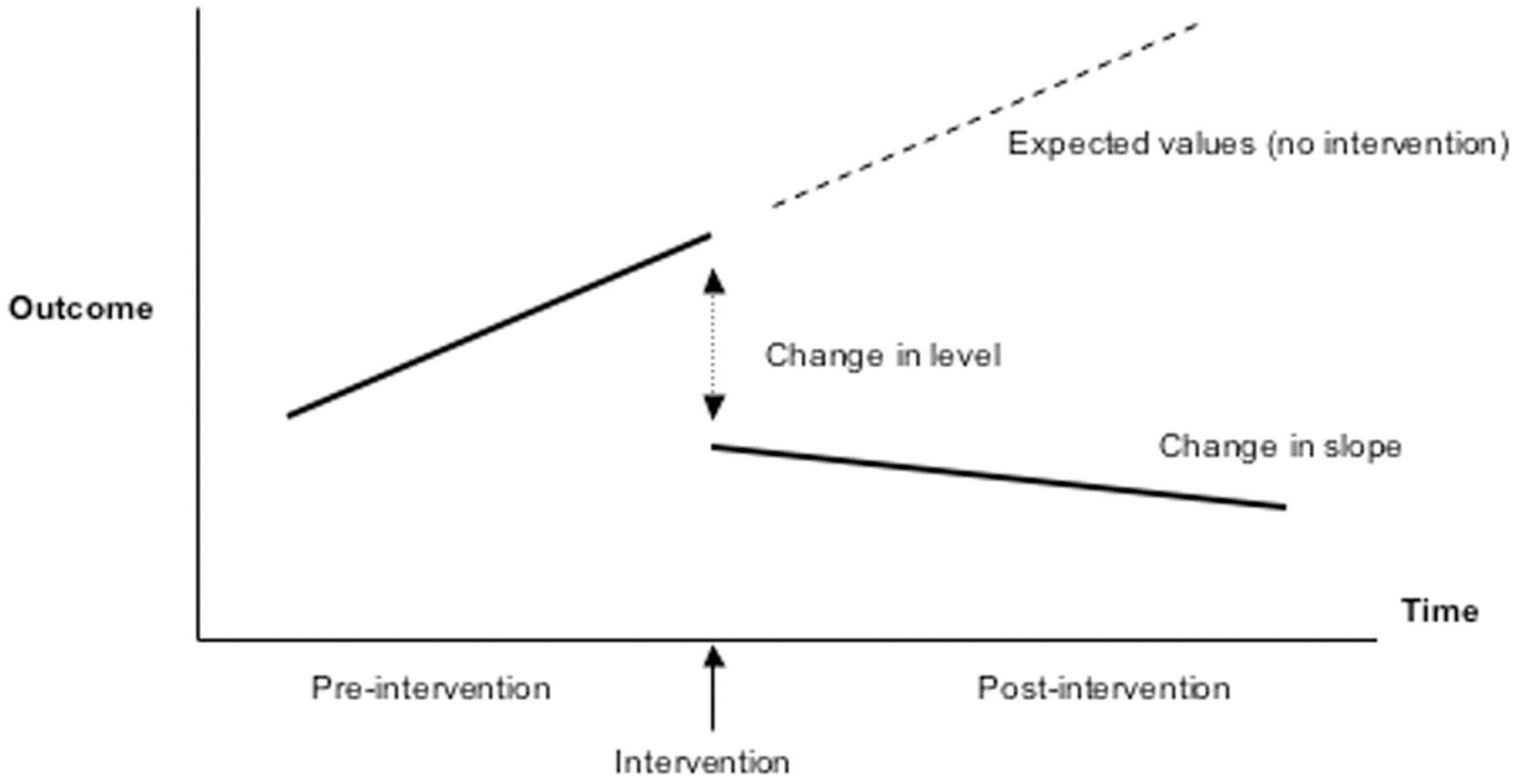


FIGURE 10.5 Percentage of students who are on-task at 10 minutes and 40 minutes into the class period. The figure presented here depicts the results of one of five classrooms investigated by Mayer et al. Only one classroom is presented here to illustrate a time-series design, whereas Mayer et al. used five classrooms and a multiple-baseline design. PLA refers to planned activity.

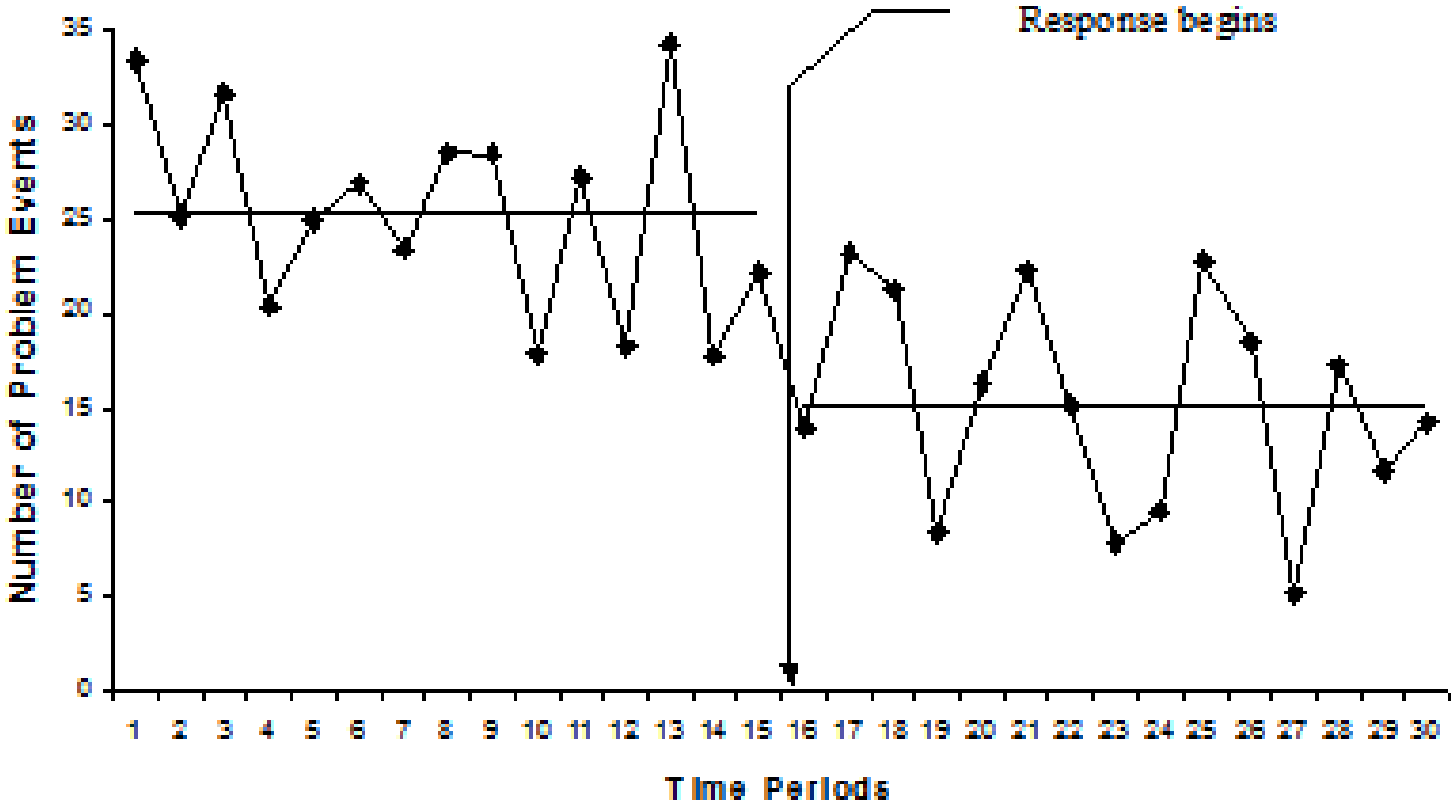
Adapted from G. R. Mayer, L. K. Mitchell, T. Clementi, E. Clement-Robertson, & R. Myatt (1993). "A dropout prevention program for at-risk high school students: Emphasizing consulting to promote positive classroom climates." *Education and Treatment of Children, 16,* 135–146. Reprinted by permission.

interrupted time series:



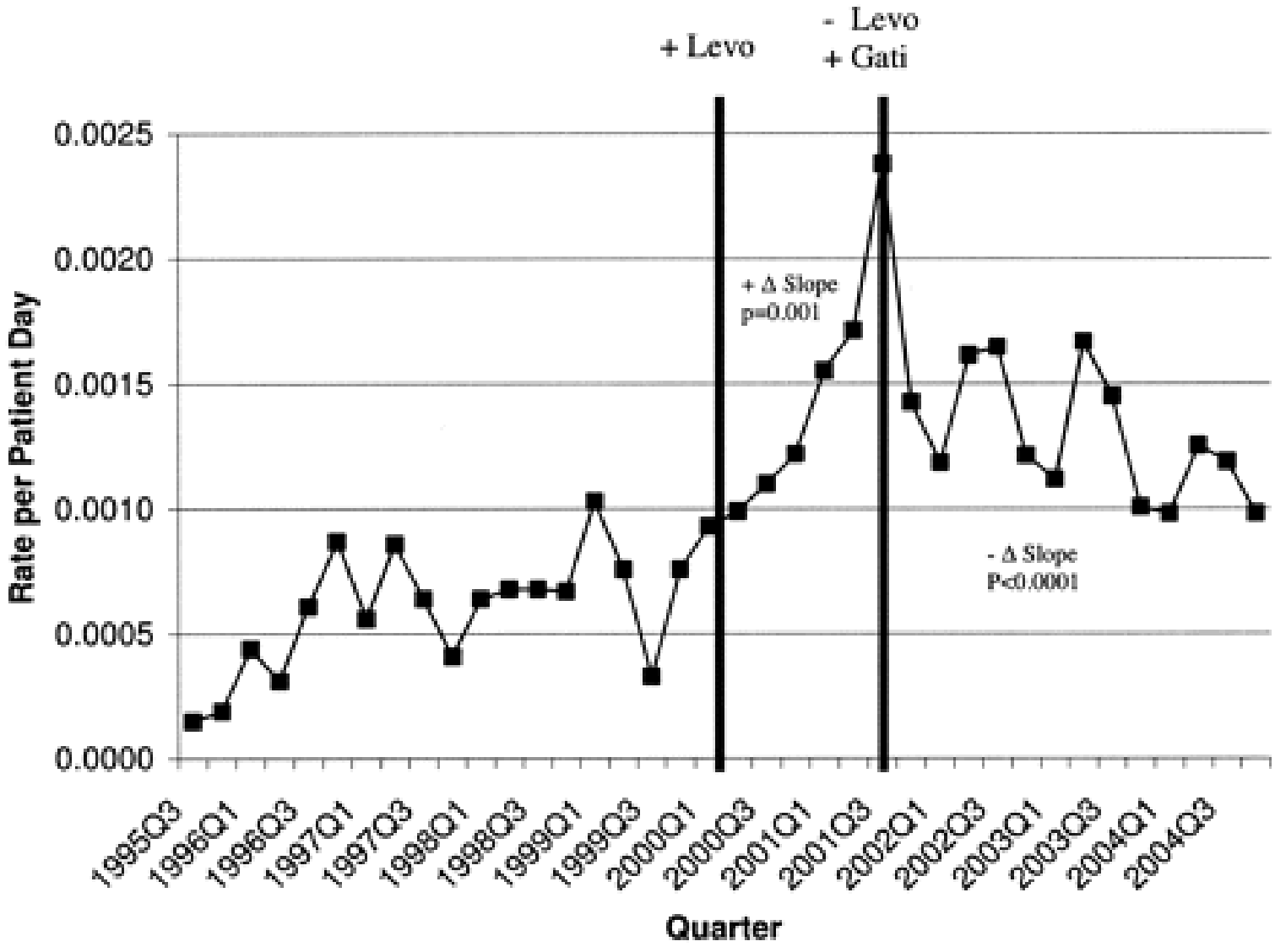
- in general look at the trend

interrupted time series:



- look at the trend: may be difficult to see response

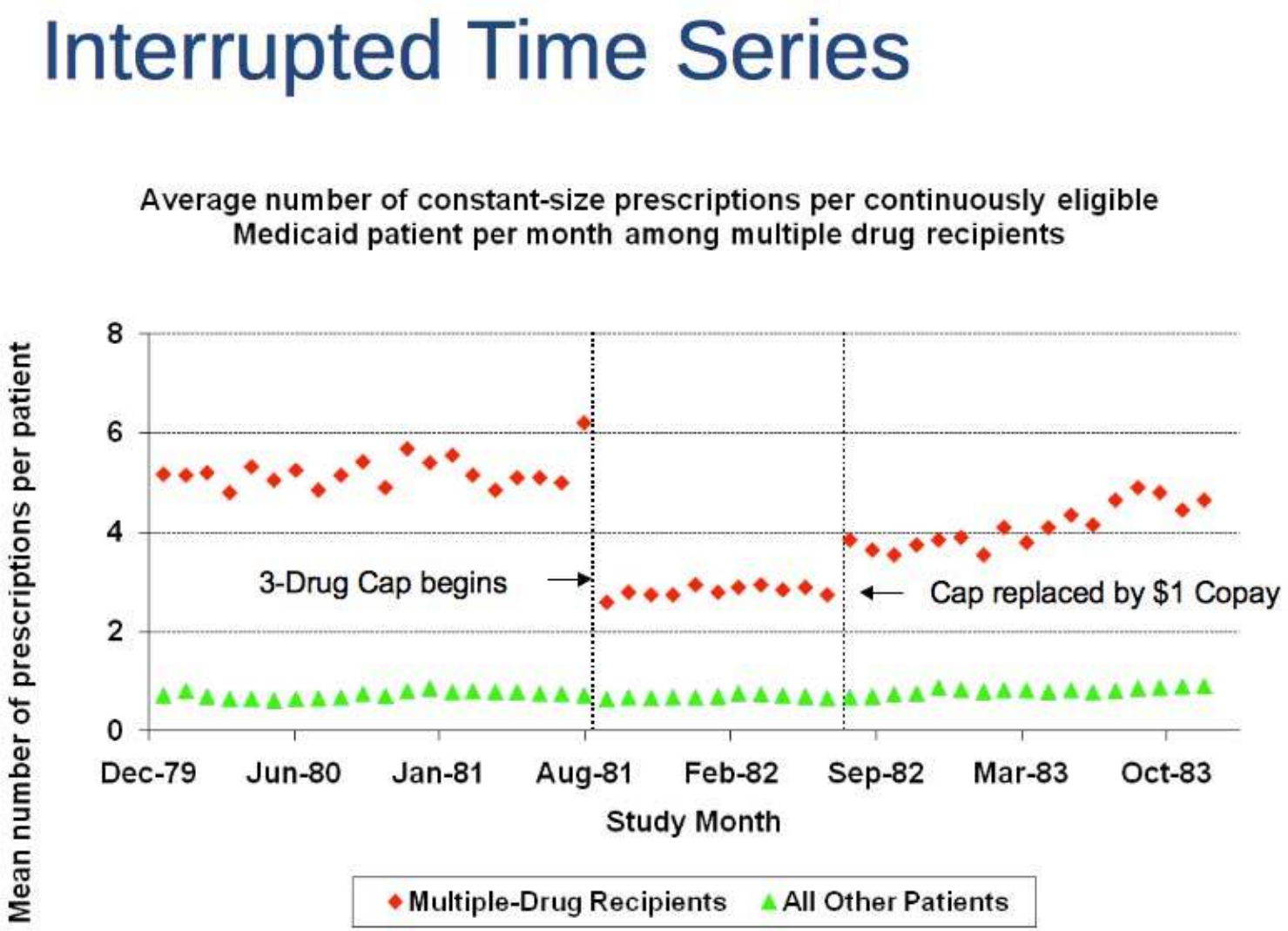
interrupted time series:



◇

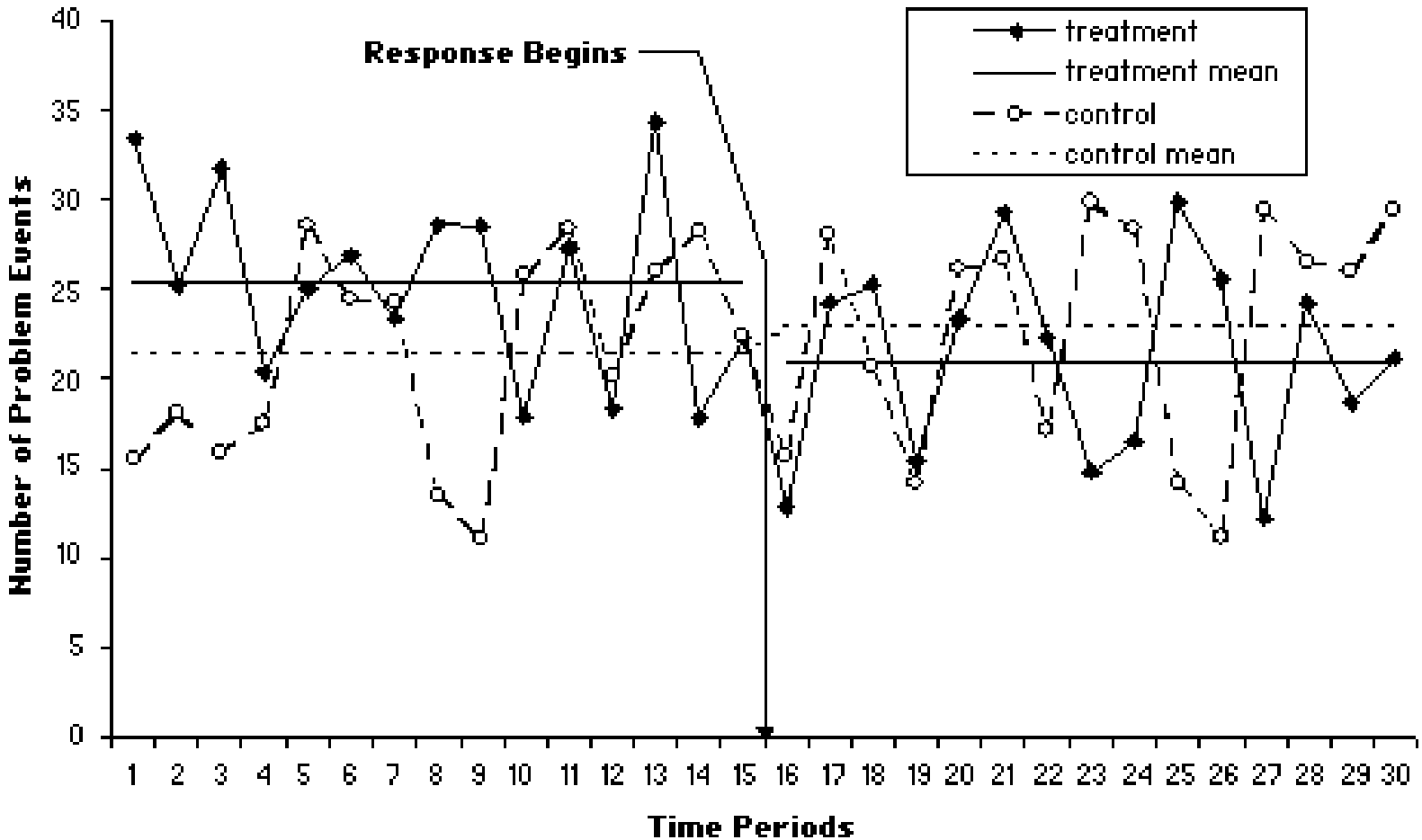
◇ more powerful: take away T→ effect dies

interrupted time series with a control



Adapted from: Soumerai et al, N Engl J Med 1987

interrupted time series with a control



◇

outline

intuition

research design basics

DID and program evaluation (Wheelan, 2013, ch13)

level of analysis

crime and policing

- remember slide from res_des.pdf about reverse causality?
- reposted next to this one for your convenience
- Wheelan (2013, p227) is giving the same example!

reverse causality OR chicken-egg dilemma

- find some exogenous (external) shock: policing$\leftrightarrow$crime
- but terror attack/alert $\rightarrow$policing$\rightarrow$crime; we know that policing$\rightarrow$crime; not the other way round

- `https://www.law.upenn.edu/fac/jklick/48JLE267.pdf`

that was natural experiment

- any other examples of natural experiments?
- also see Wheelan (p231- 2013)
- a pretty cool one is with schooling→lexp
  - natural experiment is different min school requirement
  - by state and over time

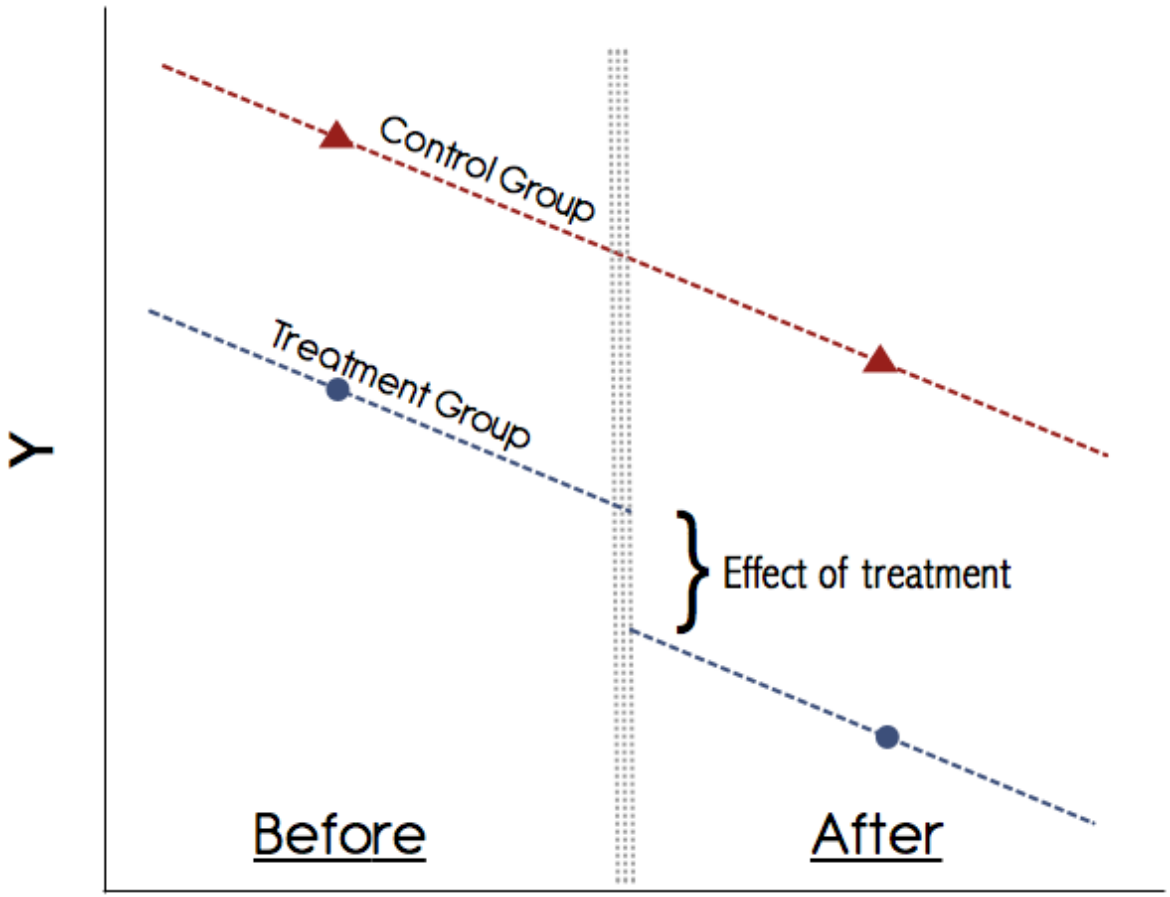
Will going to Harvard change your life?

- ◇ that's the title of ch13, and the answer is on p233
- ◇ no, it won't! [RU-Camden is just fine!]
  - (sometimes it may, though–there are always outliers)
- ◇ median salary of 130k (after 10-20yrs of exp)
  - is not only due to Harvard
  - people who go there were smart or rich or persistent
  - and either one of those can make you rich
- ◇ what counts is value added! how much univ adds
  - maybe, the worse the college, the more it actually adds!
  - people who go there may not accomplish anything
  - if not that community college or some other bad school

difference in difference (p.235 Wheelan, 2013)

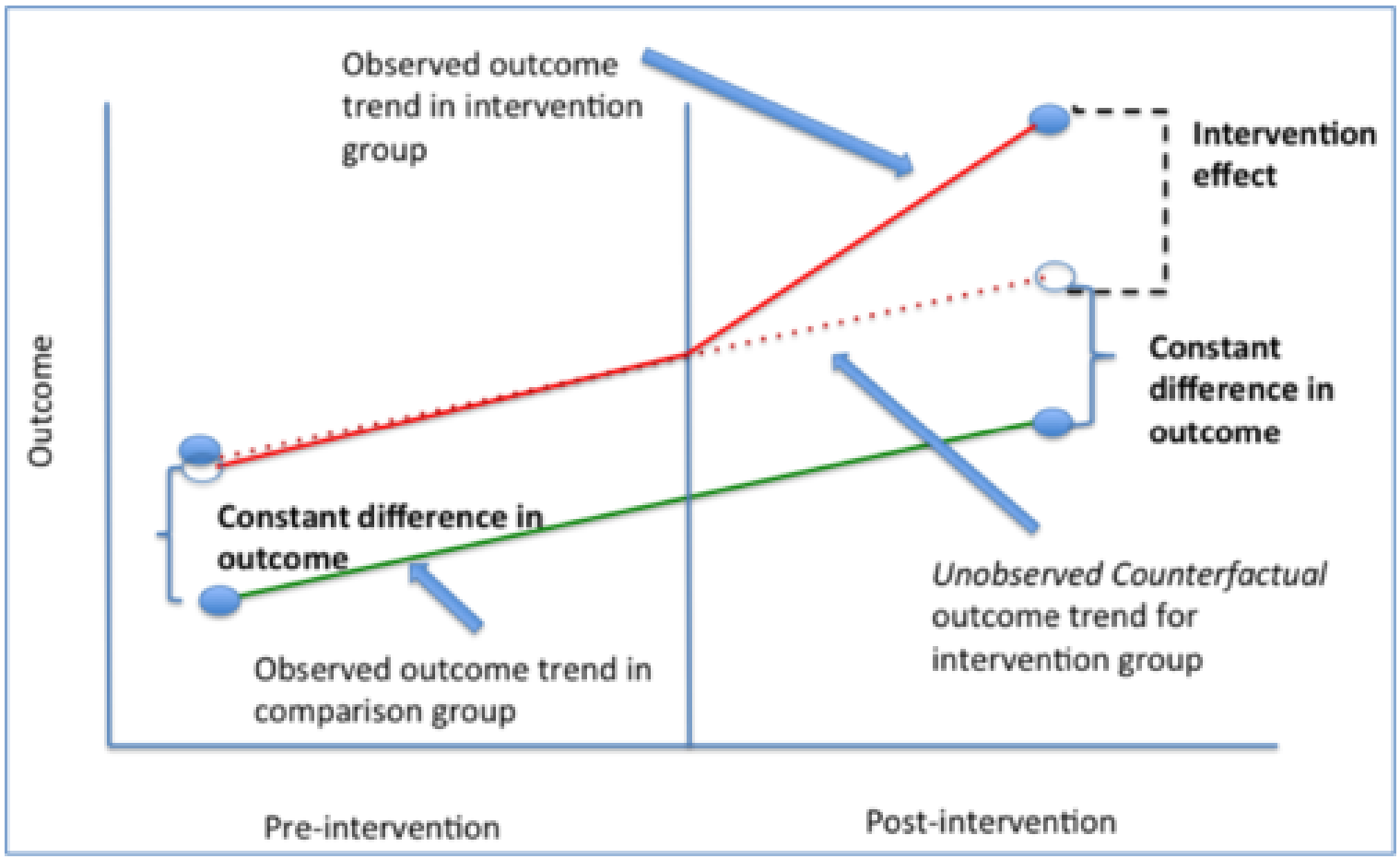
- just 'before after' with a comparison group
- did sth to one group, and not to the other group
  - over time (pre post) see if there is any difference
- like we discussed earlier in res_des.pdf
- blackboard: fig: first from p236, and then from p237
  - and pictures similar to those from res_des.pdf follow

DID


Control Group
Treatment Group
Y
Effect of treatment
Before
After

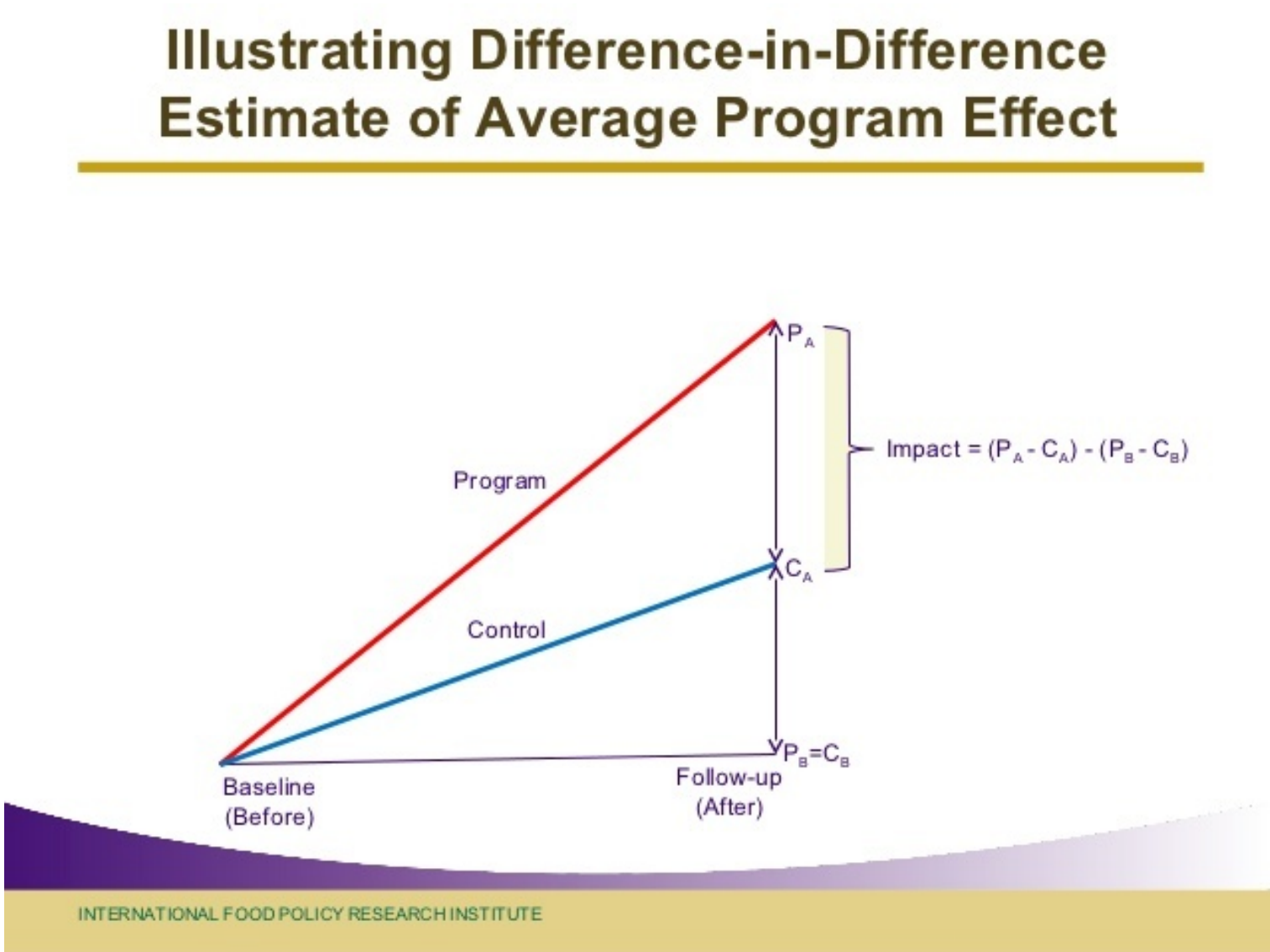
◇

DID


Observed outcome trend in intervention group
Intervention effect
Constant difference in outcome
Constant difference in outcome
Unobserved Counterfactual outcome trend for intervention group
Observed outcome trend in comparison group
Outcome
Pre-intervention
Post-intervention

◇

DID


Illustrating Difference-in-Difference Estimate of Average Program Effect
P_A
Impact = $(P_A - C_A) - (P_B - C_B)$
Program
C_A
Control
$P_B=C_B$
Baseline (Before)
Follow-up (After)
INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

◇

discontinuity analysis (p.238 Wheelan, 2013)

- ◇ can use when there is some rigid cutoff for something, say:
  - remedial program for F grades
  - prison sentence for a crime
- ◇ then compare those who just made it (C-, or a ticket)
  - v those who didn't (F, prison)–but they were just above the cutoff
- ◇ the cool thing is that the two groups are similar, especially:
  - not really any difference whatsoever with respect to cause of treatment!
  - so the treatment is arbitrary (random), so we have experiment!

example

- new jersey state government workforce profile 2010
- `http://www.nj.gov/csc/about/publications/workforce/pdf/wf2010.pdf`
- p37: minorities in state govt over time
- how increase internal validity?
- compare to PA, DE, NY etc
- factor in minority population; applications
- do experiments! many already done! again, read lit!!
  - say people with black names apply for jobs
  - students with Asian names email professors
- and both, employers and professors discriminate against!

eg: tacit knowledge is the key!

- if you know sth about state govt
  - you know that it is concentrated in Trenton
  - (one student said so)
- hence, the key is population characteristics
  - around Trenton!

next step

- ◇ if you are interested in program evaluation:
- quick `http://www.socialresearchmethods.net/kb/evaluation.php`
- in-depth, advanced: Mohr (1995), Shadish et al. (2002)

outline

intuition

research design basics

DID and program evaluation (Wheelan, 2013, ch13)

level of analysis

levels of analysis

- you are familiar with term Unit of Analysis (U/A)
- there are many levels
- there are states, counties, metropolitan areas, cities, etc
- and you often get different and even opposite conclusions depending on what level you are looking at

aggregate data

- in regional development research much of the data is aggregate
- eg income, home ownership rate at county level are sums of person-level values divided by population
- with aggregate data you are losing information
  you don't know the variability and the distribution

ecological fallacy

- it happens when you make conclusions about individual units based on group data
- eg You are on vacation in Hawaii and you meet a person from Camden, NJ
  - and you think: "she must be a criminal"
  - that Camden has the highest crime rate in the U.S. does not mean that everybody in Camden is a criminal
- now say, you meet a person that graduated from Harvard
  - and you think "she must be a genius"
  - again, just because Harvard is ranked as a best university (U.S. News) does not mean that every Harvard graduate is a genius

atomistic fallacy

- an opposite of ecological fallacy
- making inferences about groups based on individual data
- eg you found that rising individual income reduces risk of coronary heart disease (eg people stress out that they are relatively poor, they are missing out...)
- but it does not mean that increasing incomes of states would decrease coronary disease rate for a state ...
- `http://www.paho.org/english/dd/ais/be_v24n3-multilevel.htm`

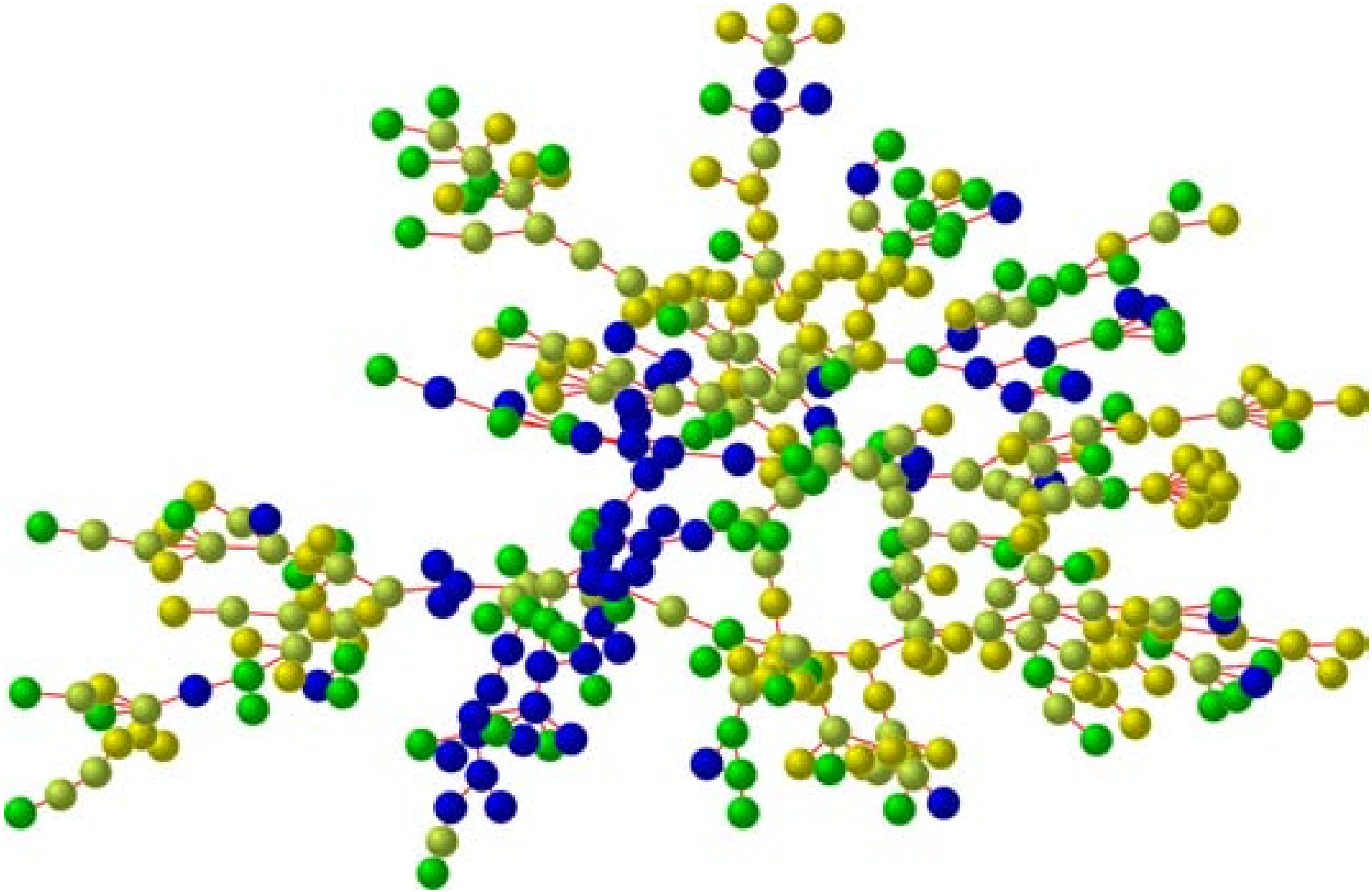
different levels, different effects

- variables at different levels may have opposite effects
- eg if i increase your salary, you'll be happier
- but if i increase salary of everybody in your county you'll be less happy
- would you like to live in a world where you make $100k and the average is $150k
- or would you like to live in a world where you make $75k and everybody and the average is $50k
- people chose the second scenario
- "a rich guy is a one who makes $100 more than his wife's sister's husband"

contextual effects

- a closely related concept is of contextual effects
- whatever you study it takes place somewhere and place matters
- so it is not only characteristics of the U/A that predict your outcome
- but also the context (characteristics of larger units in which U/A is nested)
- student is nested within a classroom, a classroom within school, a school within a district, etc etc
- a firm is nested within a city/metropolitan area/town, which is nested within a state, which is nested within a country

happiness is contagious (Fowler and Christakis, 2008)



your research project

- you should address some of the above issues in your research project
- again, a useful thing to do is be devil's advocate
  - ask yourself how/why what you are saying is not true
  - think about alternative explanations
  - what are the limitations of your study

- end every class discussing what we covered and quick look at next week
- end with a review Q&A,
- give some examples (essp in pub pol and pub adm) for concepts covered
- students will discuss concepts from the class

- quick look at next class

Fowler, J. H. and N. A. Christakis (2008): "Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis Over 20 Years in the Framingham Heart Study," British Medical Journal, Vol. 3, January 09.

Mackie, J. (1980): The cement of the universe, Clarendon Press Oxford.

Mazur, A. (2011): "Does increasing energy or electricity consumption improve quality of life in industrial nations?" Energy Policy, 39, 2568–2572.

Mohr, L. B. (1995): Impact Analysis for Program Evaluation, Sage, Beverly Hills CA, second edition ed.

Shadish, W. R., T. D. Cook, and D. T. Campbell (2002): Experimental and quasi-experimental designs for generalized causal inference, Wadsworth Cengage learning.

Wheelan, C. (2013): Naked statistics: stripping the dread from the data, WW Norton & Company.