PMAP 8131 Applied Research Methods II

Causality

Matteo Zullo

Georgia State University & Georgia Institute of Technology

June 6, 2022

Outline

Table of Contents

Epistemology

- Popper's falsification principle
- Kuhn's paradigm shift
- Lakatos research paradigm

Epistemology

Popper's falsification principle

- Logical asymmetry of verification and falsification

Epistemology

Modus ponens

- Premise I: *All dogs bark*.
- Premise II: *Tyson is a dog*.
- Conclusion: *Tyson barks*.

Modus tollens

- Premise I: *All dogs bark*.
- Premise II: *Tyson doesn't bark*.
- Conclusion: *Tyson is not a dog*.

Epistemology

Modus ponens

1. $D \rightarrow B$
2. D
3. B = Implication: 1,2

Epistemology

Modus tollens

1. $D \rightarrow B$
2. $\neg B$
3. $\neg D$ = Implication: 1,2

Epistemology

- Logical asymmetry of verification and falsification
  - One negative obs is *enough* to invalidate a theory
  - No number of positive obs would *ever* validate a theory
- Demarcationist stance
  - Falsification *demarcates* science from pseudo-science
  - Difference between statements is *logical*, not empirical

Epistemology

Duhem-Quine Thesis

1. $(T1 \wedge T2 \wedge T3) \rightarrow O$
2. $\neg O$
3. $\neg(T1 \vee T2 \vee T3)$

Epistemology

Kuhn's scientific revolutions: Popper + Duhem-Quine

- Local vs global confutations
  - Theories are bundles of propositions $T = \{T_1, \ldots, T_n\}$
  - Local confutations do not prove a theory false
- Normal science vs revolutionary science
  - Normal science is *puzzle-solving* of local confutations and never *discovers* anything new

Epistemology

Lakatos' research paradigms: Marx + Hegel

- Science as "struggle of theories"
  - Theories: $T^1 = \{T_1^1, \ldots, T_n^1\}, \ldots, T^n = \{T_1^n, \ldots, T_n^n\}$
  - A *progressive/regressive* theory generates *more/less* successes, on average, than competitor theories
- Theories as "paradigms"
  - Sub-theories layer up onto theories, only *upper layer* theories brought to the fore of evidence

Epistemology

Falsification: Regression example

- **Hypothesis**. Male-female math achievement gap due to institutional features (X_i), not genetics
- **Experimentum crucis**. Sequential regression models:

$$MATH_i = \beta_0 + \beta_1 Female_i + \epsilon_1 \quad (1)$$

$$MATH_i = \beta_0 + \beta_1 Female_i + \beta_2 X_i + \epsilon_1 \quad (2)$$

Epistemology

Falsification: Regression example

- **Hypothesis**. Male-female math achievement gap due to institutional features (X_i), not genetics
- **Experimentum crucis**. Sequential regression models:

$$MATH_i = \beta_0 + \beta_1 Female_i + \epsilon_1 \quad (1)$$
$$MATH_i = \beta_0 + \beta_1 Female_i + \beta_2 X_i + \epsilon_1 \quad (2)$$

- **Decision rule**: Reject hypothesis if β_1 significant in model 1), but insignificant in model 2)

Epistemology

Falsification: Regression example (Guiso et al., 2008)

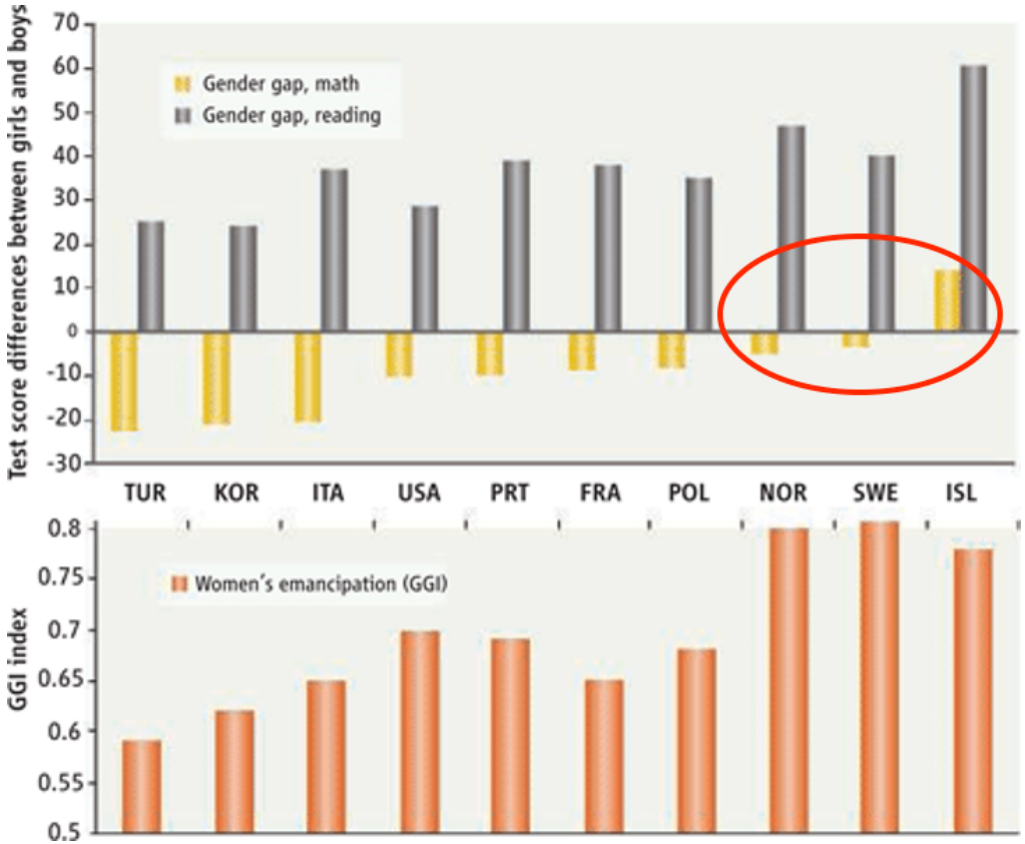


Figure: Female-male gap reverts in countries with greater gender parity

Epistemology

Human subjects

- *Act in such a way that you treat humanity, whether in your own person or in the person of another,* **always at the same time as an end** *and never simply as a means* (Immanuel Kant)

Human subjects: Violations

- Violence on experimentees
- Framing of questions
- Omission of information
- Interaction with experimentees

Epistemology

Data management

- *There are three kinds of lies: Lies, Damned Lies,* **and Statistics** (Mark Twain)

Data management: Violations

- Addition or removal of observations
- "P-hacking"
- Framing of findings
- Uncommented code!!!

Table of Contents

Causation

Causation

1. Correlation
2. Time-order
3. Non-spuriousness

Causation

1. Correlation

$$\rho(X,Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \approx \frac{COV_{XY}}{SD_X SD_Y} \approx \frac{\sum(X-\overline{X})(Y-\overline{Y})}{SD_X SD_Y}$$

- Correlation: Examples
  - Size of left and right foot
  - Watching football and earnings
  - Leap year and Olympic Games

Causation

2. Non-spuriousness
   - Relationship between X and Y spurious to Z
- Non-spuriousness: Examples
  - Air conditioning and IQ test score (Z: Temperature)
  - Suicide and Reddit engagement (Z: Depression)
  - Car usage and individual longevity (Z: Income)

Causation

3. Time-order
   - X comes before Y (Beware: reverse causation!)

Table of Contents

Validity

Experiment

- Subject features
- Experimental conditions
- Experimental stimulus

Validity

- Internal validity
  - Equality of treated and controls
- External validity
  - Equality of experimentees and outer population

Validity

- Internal validity
  - Equality of treated and controls
    - Observables and unobservables features of treatments that affect the outcome match features of controls
- External validity
  - Equality of experimentees and outer population
    - Experimental conditions and stimulus match inputs in the real world

Threats to Validity

- Threats to internal validity
  - History
  - Maturation
  - Instrumentation
  - Regression to the mean
  - Self-selection
  - Differential mortality

Threats to Validity

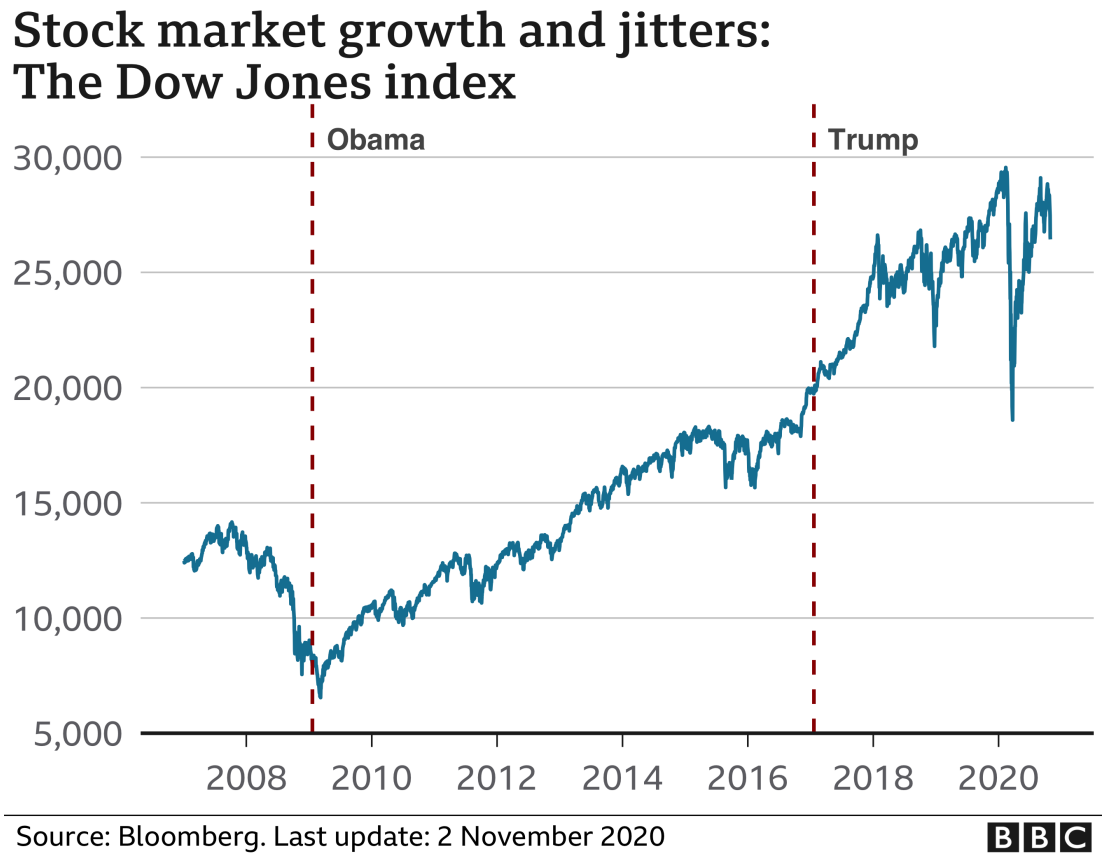
- Threats to external validity
  - Measuring
  - Testing
  - General equilibrium effects
  - Multiple treatments
  - Self-selection

Threats to Validity

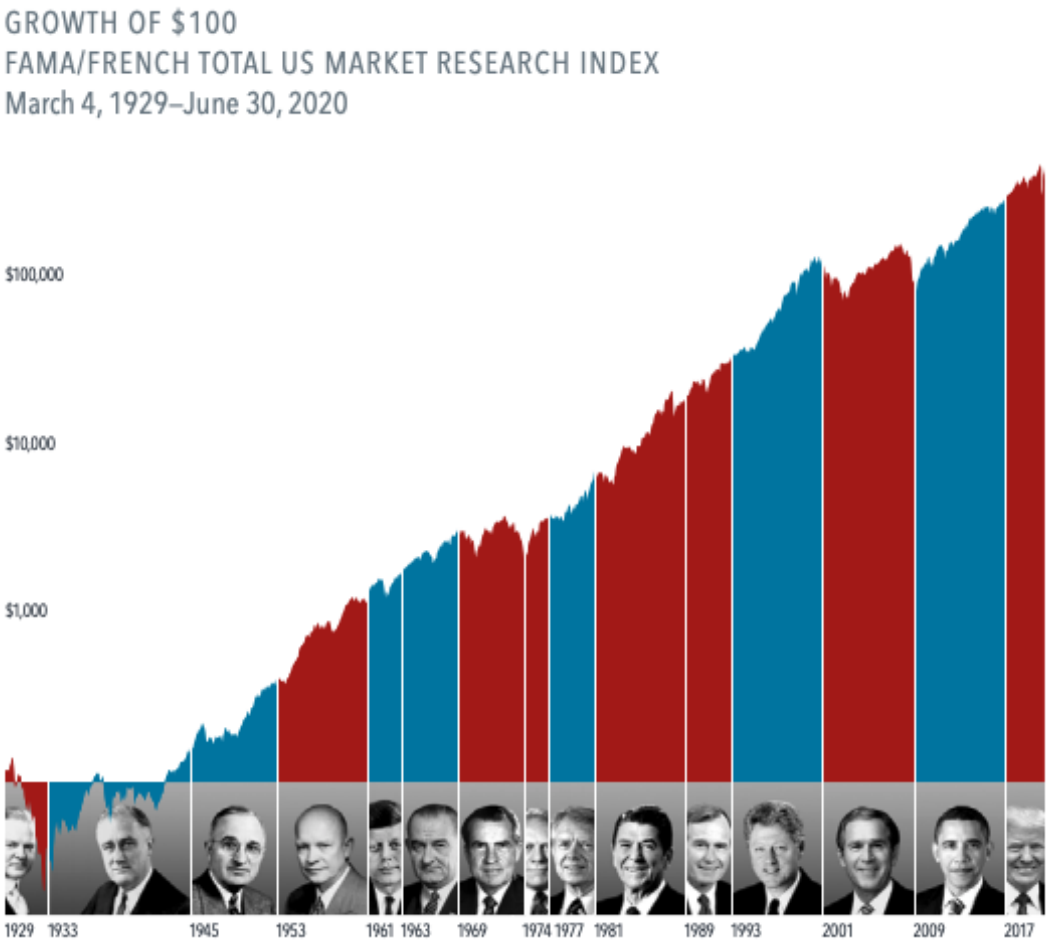
History. *Treatment effect are attributable to the passing of time*.

- Example: Federal policies and economic growth rates
  - Economy follows own short- and long-run trends
  - Political cycle overlaps with economy, does not cause it

Threats to Validity



Threats to Validity

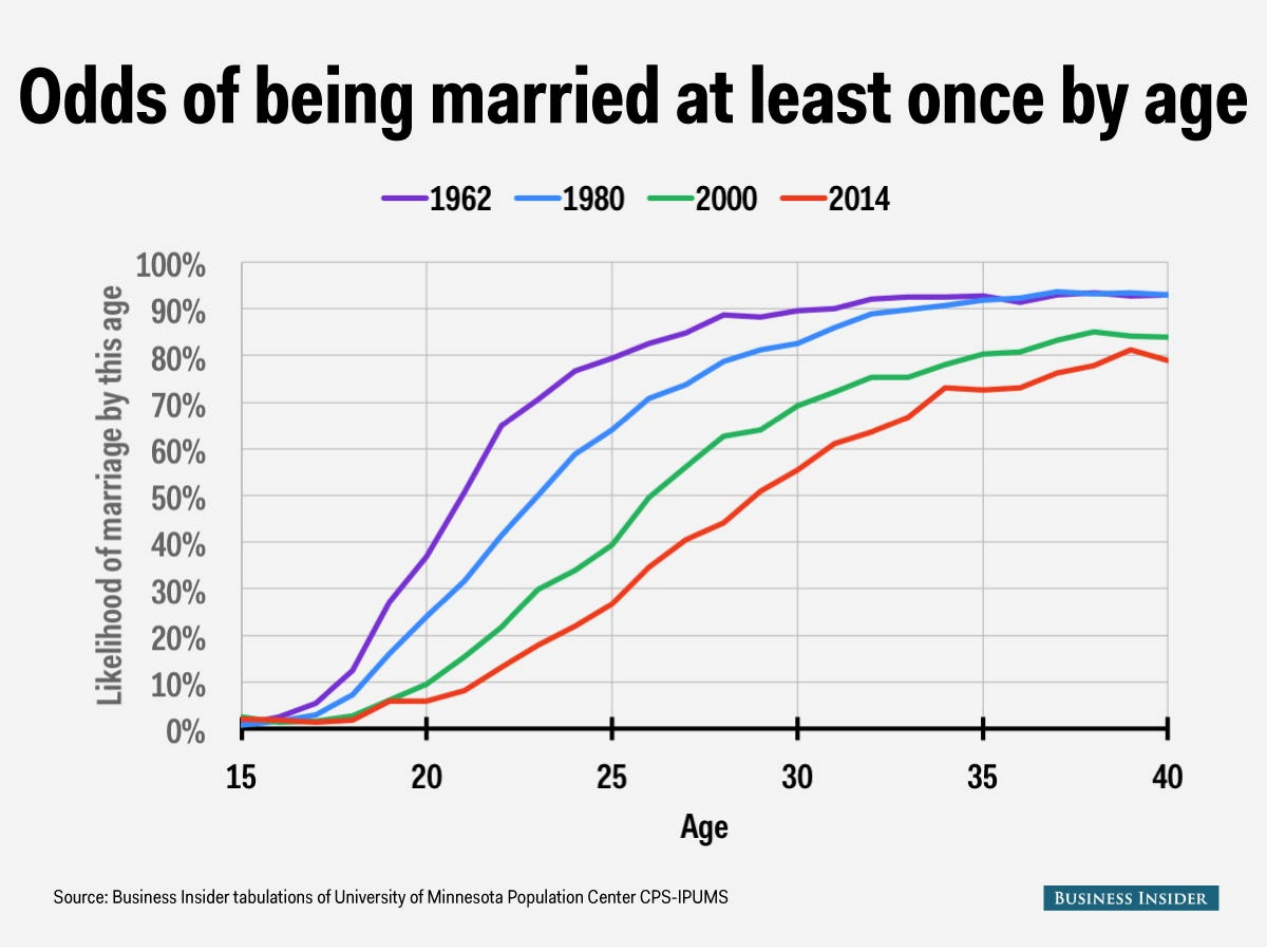


Threats to Validity

Maturation. *Treatment effect are attributable to the aging of subjects*.

- Example: Welfare programs and marital outcomes
  - People more likely to marry and have kids as they age

Threats to Validity



Threats to Validity

Instrumentation. *The effect is measured with instrumental and nonrandom error*.

- Example: Domain knowledge tests
  - Test measures test-taking ability, not domain knowledge

Threats to Validity

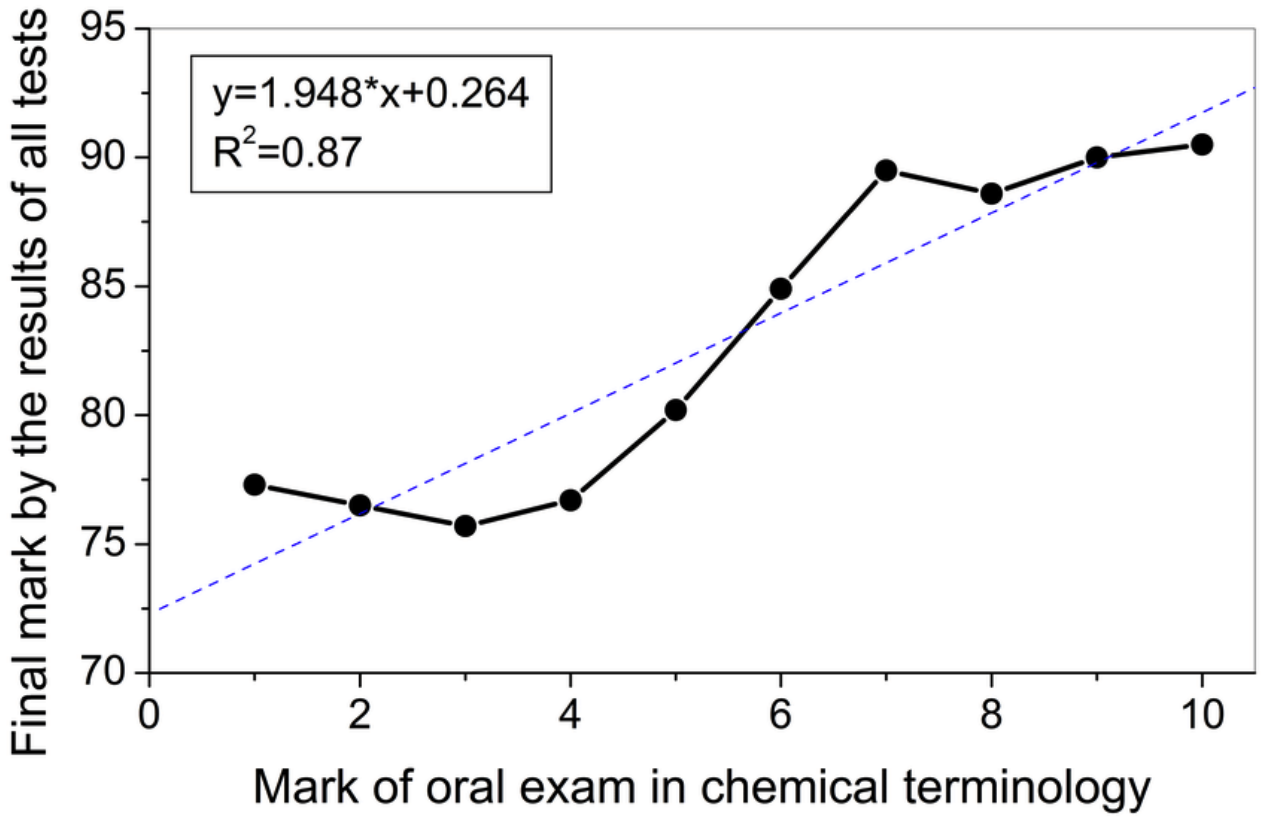


Figure: Correlation of chemistry test scores and other tests (Gryshchenko et al., 2021)

Threats to Validity

Regression to the mean. *Treatment is observed at a point below or above their average value*.

- Example: Job training programs
  - Trainees likely to get back to workforce anyways

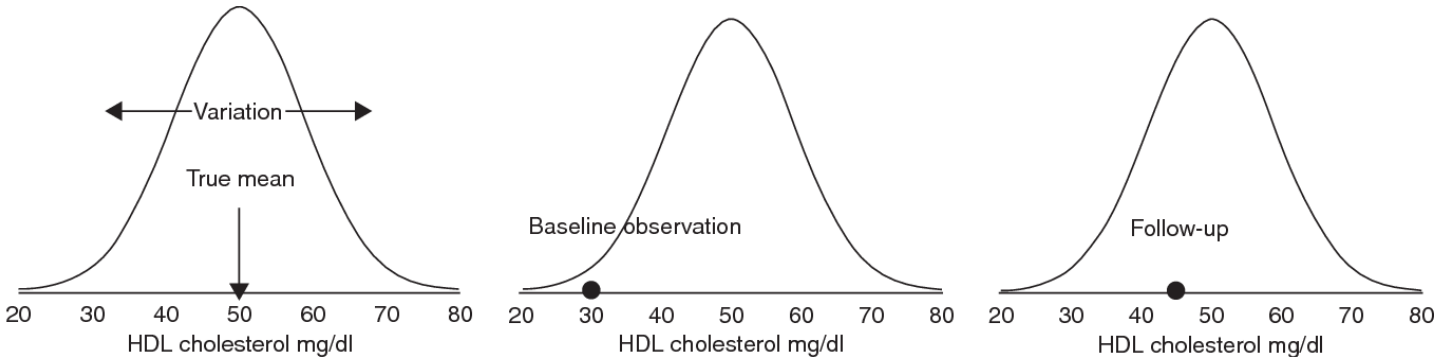


Figure: Regression to baseline cholesterol (Barnett et al., 2005)

Threats to Validity

Self-selection. *Features of experiment drive selection into treatment group*.

- Example: Immigrant earnings
  - Migrants systematically different from those who stay

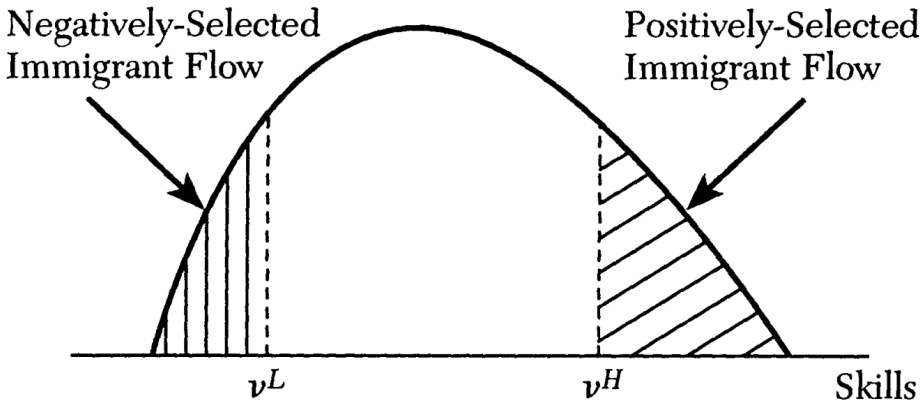


Figure: Immigrant self-selection (Borjas, 1994)

Threats to Validity

Differential mortality. *Treatment and control exit experiment at different and nonrandom rates*.

- Example: Welfare programs
  - Burdensome paperwork forfeits unmotivated treatments

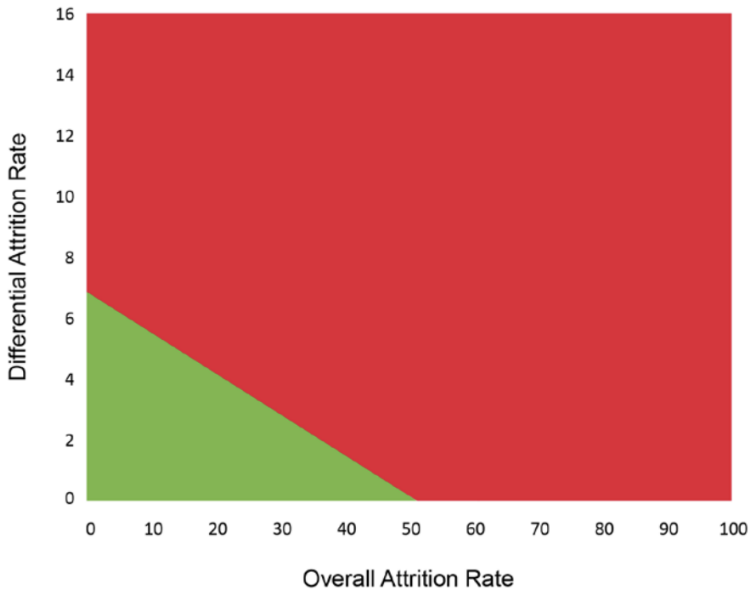


Figure: Maximum differential attrition (Mathematica, 2015)

Threats to Validity

Measuring. *Reaction to measuring affects the outcome*.

- Example: Observing experimentees
  - Behaviors change when we know we are being measured

Threats to Validity

Testing. *Reaction to testing affects the outcome*.

- Example: Pre-testing
  - Pre-testing increases test-taking skills

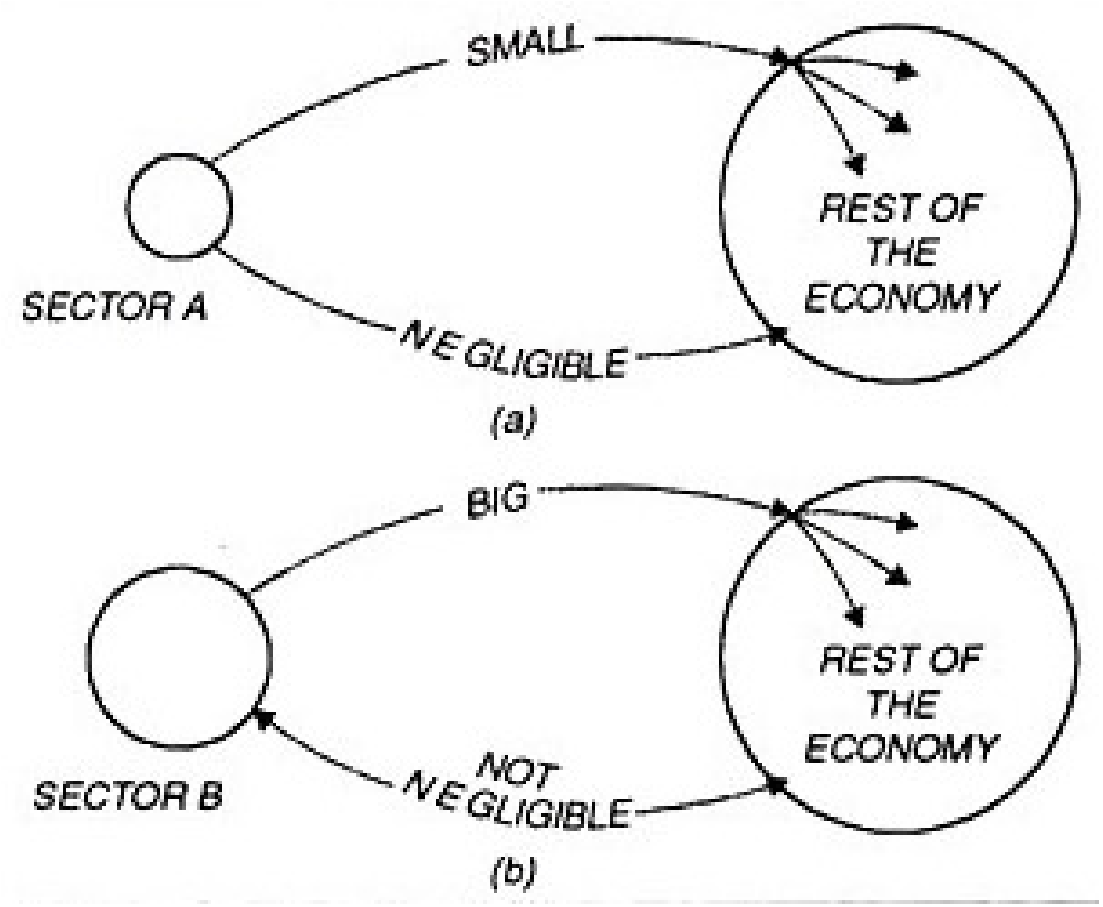
Group	*Pretest*	*Treatment*	*Posttest*
1	O_1	X	O_2
2	O_1	—	O_2
	Time ▶		

Threats to Validity

General equilibrium effects. *Generalization of treatment to population forfeits local effects*.

- Example I: The "paradox of schooling"
  - If I get an A, I stand out. If everyone gets an A, no one stands out
- Example II: Government spending
  - If I get a subsidy, I am better off. If everyone gets a subsidy, inflation makes us all worse off

Threats to Validity



Threats to Validity

Multiple treatments. *Treatment effect does not generalize if treatment is administered jointly with a second treatment*.

- Example: Psychological counseling
  - If counseling involves commutes, treatment effect is the joint effect of counseling and taking time off work

Threats to Validity

Self-selection. *Features of experiment drive selection into experimental group*.

- Example: Student evaluations of professors
  - Students leave evaluations when very happy or unhappy