

Randomized Evaluations: Design & Ethics

Erick Gong

UC Berkeley & CEGA

September 8

Agenda

- Randomized Evaluations (Design & Ethics)
- Dupas (2009)
- Group Project
- Data

HIV Prevention Programs



Ethical Concerns

- What are major ethical concerns for implementing a randomized evaluation?

Ethical Concerns

- What are major ethical concerns for implementing a randomized evaluation?
 - 1 Is the R.E. Fair (Resource Allocation)?
 - 2 Does evaluation harm anyone?
(Withholding intervention that we know benefits them)
(Unintended Consequences of intervention)

Fairness

- Constrained Resources

Fairness

- Constrained Resources

Budget Constraints would have prevented everyone from receiving intervention initially.

Fairness

- Constrained Resources

Budget Constraints would have prevented everyone from receiving intervention initially.

- Phase In Roll Out

Everyone eventually receives intervention (control group is the group left out initially)

This allows us to also see if the intervention has an effect (Does it work?)

Fairness

- Constrained Resources

Budget Constraints would have prevented everyone from receiving intervention initially.

- Phase In Roll Out

Everyone eventually receives intervention (control group is the group left out initially)

This allows us to also see if the intervention has an effect (Does it work?)

- Randomization can be transparent

Reduces misapplication due to favoritism or corruption

Fairness

- Constrained Resources

Budget Constraints would have prevented everyone from receiving intervention initially.

- Phase In Roll Out

Everyone eventually receives intervention (control group is the group left out initially)

This allows us to also see if the intervention has an effect (Does it work?)

- Randomization can be transparent

Reduces misapplication due to favoritism or corruption

- What are complaints against it?

Fairness

- Constrained Resources

Budget Constraints would have prevented everyone from receiving intervention initially.

- Phase In Roll Out

Everyone eventually receives intervention (control group is the group left out initially)

This allows us to also see if the intervention has an effect (Does it work?)

- Randomization can be transparent

Reduces misapplication due to favoritism or corruption

- What are complaints against it?

Intervention might not be focused on population deemed most in need
However, you can always randomly assign intervention within “most in need” group

Hippocratic Oath

First Do No Harm

Ethics & HIV

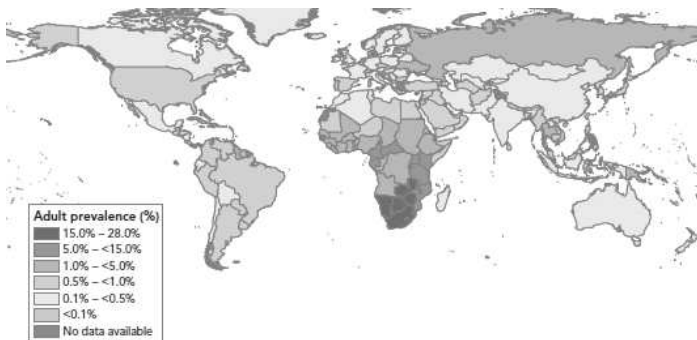
- Why would interventions in HIV prevention be different from other interventions?

Ethics & HIV

- Why would interventions in HIV prevention be different from other interventions?

There is a great cost of being incorrect. Interventions that increase the risk of HIV may have *irreversible* effects.

HIV/AIDS in SSA



HIV/AIDS Prevention

What are ways of reducing new HIV infections?

HIV/AIDS Prevention

What are ways of reducing new HIV infections?

- Biomedical methods: vaccine, male circumcision
- Behavioral Change:
 - Reducing number of partners
 - Increasing condom usage during sexual intercourse

Ethics Research

- Why are evaluations of interventions in HIV prevention of particular ethical concern?
- What are the effects of an intervention?
- Difficult to fully understand all the unintended consequences

Dupas (2009)

HIV Information to Kenyan Teenagers

- Does providing information about HIV prevention to Kenyan teenagers affect behavior?
- What sort of information would you provide?

Extensive vs. Intensive Margin

- Two types of messages (information) you can give:
- Extensive Margin
Do I have sex or not?
- Intensive Margin
If I have sex, who do I choose to have sex with?
- Which message do you think would have larger impact
“Don’t have sex until you get married”
“Certain types of people might have higher likelihood of being infected”

Theory

Adult Men differ from Teenage Boys

- Higher HIV Prevalence Rate
- Greater Financial Resources
- Derive relatively less utility from condom-protected sex

Girls choose sexual partners based on:

- Perceived riskiness of partnership
- Size of compensation transfer

Trade off between risk (HIV infection) and compensation?

Dupas (2009)

HIV information for Kenyan teenagers

- National HIV Prevention Curriculum:
- Encourage Abstinence Until Marriage
- Average HIV prevalence in the population
- Teacher reinforcement training

NGO Relative Risks Campaign:

- Information on HIV prevalence disaggregated by gender and age group
- Presentation by trained NGO officer

Experiment

Provide Relative Risks Information in a randomly selected subset of schools

- Treatment Cohort: 13,000 8th graders at time of campaign (2004)
- Treatment Students: 2,500 Students enrolled in treatment schools during campaign
- Comparison Cohort: Students a year ahead of or below treatment cohort

Experiment (2)

Treatment Arm 1

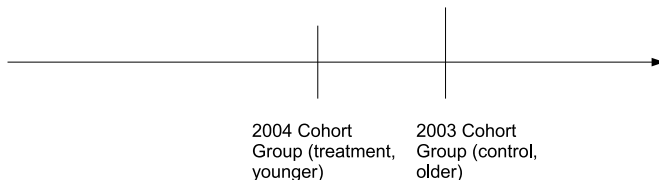
- Teacher Training on HIV/AIDS curriculum
-

Treatment Arm 2

- Survey
- 10 minute video on sugar daddies
- Discussion on Cross Generational Sex
- Distribution of HIV Prevalence

Control Arm

- Schools that did not receive the treatment
- Younger cohort group in same school as treatment but not exposed to relative risk information



Outcomes

- What is the outcome we want to measure?

Outcomes

- What is the outcome we want to measure?

Risky Sexual Behavior (Unprotected Sex, Sex with partners with a higher likelihood of being HIV)

- What can we use instead?

Outcomes

- What is the outcome we want to measure?

Risky Sexual Behavior (Unprotected Sex, Sex with partners with a higher likelihood of being HIV)

- What can we use instead?

Pregnancies

Self-reported Sexual Behavior

- Problems with these two outcomes?

Outcomes

- What is the outcome we want to measure?

Risky Sexual Behavior (Unprotected Sex, Sex with partners with a higher likelihood of being HIV)

- What can we use instead?

Pregnancies

Self-reported Sexual Behavior

- Problems with these two outcomes?

Pregnancies (Abortions, Birth-Control)

Self reported sexual behavior (bias)

Data Collection

- Collection of Pregnancy Data

At each visit, the list of all students on the 2004 enrollment form was read aloud to pupils enrolled in upper grades in 2005, and for each of the students on the list, the following questions were asked: Is X still in school? If yes, in what grade? In what school? Does she still live in the area? Is she married? Does she have any children? If so, how many? How old is her first born? Is she pregnant?

- Concerns?

Data Collection

- Collection of Pregnancy Data

At each visit, the list of all students on the 2004 enrollment form was read aloud to pupils enrolled in upper grades in 2005, and for each of the students on the list, the following questions were asked: Is X still in school? If yes, in what grade? In what school? Does she still live in the area? Is she married? Does she have any children? If so, how many? How old is her first born? Is she pregnant?

- Concerns?

Responses to questions verbal or written?

List of female students read aloud?

Data Collection

- Collection of Pregnancy Data

At each visit, the list of all students on the 2004 enrollment form was read aloud to pupils enrolled in upper grades in 2005, and for each of the students on the list, the following questions were asked: Is X still in school? If yes, in what grade? In what school? Does she still live in the area? Is she married? Does she have any children? If so, how many? How old is her first born? Is she pregnant?

- Concerns?

Responses to questions verbal or written?

List of female students read aloud?

How would you alleviate these concerns?

Intervention: Relative Risks

Age	15-19	20-24	25-29	30-39
Female	22%	36%	35%	32%
Male	4%	13%	28%	32%

Information Misuse

- Will this information lead to safe behavior?
- Can this information be misused?

Information Misuse

- Will this information lead to safe behavior?
- Can this information be misused?

Teenage girls may engage in more unprotected sex with male teenagers

Older males may get information and seek young girls for unprotected sex

Empirical Strategy

Compare outcomes (Y) one year later of students in treatment (relative risk) vs. control groups

- Self-reported sexual behavior
- Childbearing and marital status
- Characteristics of sexual partners of girls who had started childbearing

What are the two control groups?

Empirical Strategy

Compare outcomes (Y) one year later of students in treatment (relative risk) vs. control groups

- Self-reported sexual behavior
- Childbearing and marital status
- Characteristics of sexual partners of girls who had started childbearing

What are the two control groups?

Control Schools

Older Cohort Group

Results

SPECIFICATION MODEL	Has started childbearing		
	SD	SD	DD
	OLS	PROBIT (Δ ME)	OLS
	(1)	(2)	(3)
RR Information	-0.015 (0.008)**	-0.014 (0.007)**	0.004 (0.011)
RR Information x 2004 Cohort			-0.024 (0.013)*
TT on HIV/AIDS Curriculum	0.006 (0.006)	0.008 (0.006)	0.000 (0.005)
<u>Sample</u>			
Control Cohort Included (2003 cohort)			Yes
<u>Controls</u>			
Individual Characteristics	Yes	Yes	Yes
Primary School Characteristics	Yes	Yes	Yes
Primary School Fixed Effects			
Observations	5989	5989	10970
Mean of Dep Var (RR=0)	0.054	0.054	0.054

Results 2

SPECIFICATION MODEL	Age Difference between Teenage Girl and her partner	
	SD	DD
	OLS (1)	OLS (2)
RR Information	-1.734 (0.571)***	1.035 (0.858)
RR Information x 2004 Cohort		-2.548 (1.061)**
TT on HIV/AIDS curriculum	-0.811 (0.711)	-0.409 (0.472)
<u>Sample</u>		
Control Cohort Included (2005 cohort)		Yes
<u>Controls</u>		
Individual characteristics	Yes	Yes
Primary School Characteristics	Yes	Yes
Observations	120	250
Mean of Dep Var (RR=0)	5.84	5.84
Std. Dev.	(4.21)	(4.21)

Increases Intensive Margin

	Ever had sex	
	OLS	PROBIT
		(ME)
RR Information	0.101 (0.031)+++	0.090 (0.026)+++
TT on HIV/AIDS curriculum	-0.028 (0.023)	-0.027 (0.022)
Observations	2173	2173
Mean of Dep Var (RR=0)	0.191	0.191
<u>Controls</u>		
Individual Characteristics	Yes	Yes
Secondary School Characteristics	Yes	Yes

Findings

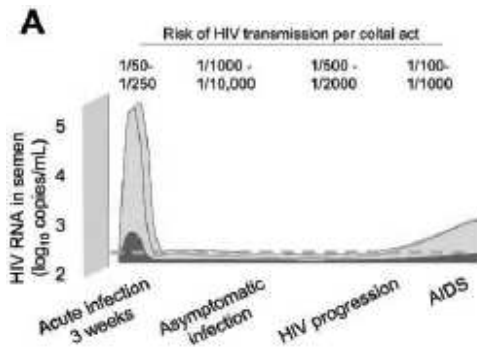
- Reduction in teenage pregnancy (indicator of safer sex)
- Those who do get pregnant, have partners of lower age
- Increases in sexual behavior between young women and young men
- No changes in extensive margin “No sex” but changes in intensive margin “partner selection”
- Any ethical concerns?

HIV transmission

- What is the transmission rate of HIV?

HIV transmission

- What is the transmission rate of HIV?



HIV Transmission Risk

- Most at risk when having unprotected sex with partner during acute infection phase (3 weeks)
- If teenage girls substitute from older partners to younger partners, under certain scenarios, could be at *increased* risk of infection

Ethics (1)

- If we give people information on HIV, they respond with their behavior
- Unethical to withhold information that we know is right, if it can be implemented.

but

- Unethical to give out information that we are not sure is right, or may be misleading

Ethics (2)

- Deny control group the intervention,

but:

- Don't actively hurt them
- Don't give wrong information
- Don't make them worse off than they would otherwise be

Can give a token gift/compensation (but small enough not to make it a treatment in itself)

Practically Speaking

- Evaluations require approval from IRB (Institutional Review Board)
- Human Subjects Protocol

Intermission

Research Proposal

- Presentation / Written Proposal
- Presentations will be in the last two weeks of class
- Groups of 4
- Do not stress

Research Proposal

New impact evaluation (ex-ante)

- Actual program that is beginning to roll out
- New intervention that you want to propose to a government agency / NGO

Retrospective Impact Evaluation (ex-post)

- Program has already been implemented
- Data is available from the program

Guidelines

- 1 Describe what you want to evaluate. Motivate why it is important. What is the population you are examining (i.e. school age children in rural Kenya, women of child bearing age in India, etc.)
- 2 What outcomes will you use to evaluate? Can you measure them?
Education: test scores, attendance, grade completion
Health: weight, height
Empowerment, Corruption (harder to do)
- 3 Technique
New impact evaluation (randomized evaluation):
Retrospective Study (diff-in-diff, regression discontinuity, propensity score matching)
Explain why you simply cannot compare people who get the treatment vs. those that don't
- 4 Data Analysis
New evaluation: power calculation, some statistics about the population (age, gender, education)
Retrospective Study: summary statistics & preliminary regressions
- 5 Policy Implications

Optional

- Externalities

If the intervention you are considering could effect those other than the treatment group, what type of effect might it have?
How might this effect your results?

Example: De-worming

Data Sources

- World Bank Living Standards Measurement Study:
go.worldbank.org/PDHZFQZ6L0
- Demographic Health Surveys: www.measuredhs.com/
- American Economic Association: www.aeaweb.org/aer/index.php
- American Economic Journal: Applied Economics:
www.aeaweb.org/aej/app/index.php
- Professor Michael Kremer (Harvard):
www.economics.harvard.edu/faculty/kremer/data_sets_kremer