Causal Inference

Mixtape Session

Roadmap

- Counterfactuals and causality
  - Causality and models
  - Potential outcomes
  - Randomization and selection bias
  - Randomization inference
- Directed Acyclic Graphs
  - Graph notation
  - Backdoor criterion
  - Collider bias
  - Front door criterion
  - Concluding remarks

Demand for Learning HIV Status

- Rebecca Thornton implemented an RCT in rural Malawi for her job market paper at Harvard in mid-2000s
- At the time, it was an article of faith that you could fight the HIV epidemic in Africa by encouraging people to get tested; but Thornton wanted to see if this was true
- She randomly assigned cash incentives to people to incentivize learning their HIV status
- Also examined whether learning changed sexual behavior.

Experimental design

- Respondents were offered a free door-to-door HIV test
- Treatment is randomized vouchers worth between zero and three dollars
- These vouchers were redeemable once they visited a nearby voluntary counseling and testing center (VCT)
- Estimates her models using OLS with controls

Why Include Control Variables?

- To evaluate experimental data, one may want to add additional controls in the multivariate regression model. So, instead of estimating the prior equation, we might estimate:

$$Y_i = \alpha + \delta D_i + \gamma X_i + \eta_i$$

- There are 2 main reasons for including additional controls in the regression models:
  1. Conditional random assignment. Sometimes randomization is done *conditional* on some observable (e.g., gender, school, districts)
  2. Exogenous controls increase precision. Although control variables X_i are uncorrelated with D_i, they may have substantial explanatory power for Y_i. Including controls thus reduces variance in the residuals which lowers the standard errors of the regression estimates.

Table: Impact of Monetary Incentives and Distance on Learning HIV Results

	1	**2**	**3**	**4**	**5**
Any incentive	0.431***	0.309***	0.219***	0.220***	0.219 ***
	(0.023)	(0.026)	(0.029)	(0.029)	(0.029)
Amount of incentive		0.091***	0.274***	0.274***	0.273***
		(0.012)	(0.036)	(0.035)	(0.036)
Amount of incentive^2			-0.063***	-0.063***	-0.063***
			(0.011)	(0.011)	(0.011)
HIV	-0.055*	-0.052	-0.05	-0.058*	-0.055*
	(0.031)	(0.032)	(0.032)	(0.031)	(0.031)
Distance (km)				-0.076***	
				(0.027)	
Distance^2				0.010**	
				(0.005)	
Controls	Yes	Yes	Yes	Yes	Yes
Sample size	2,812	2,812	2,812	2,812	2,812
Average attendance	0.69	0.69	0.69	0.69	0.69

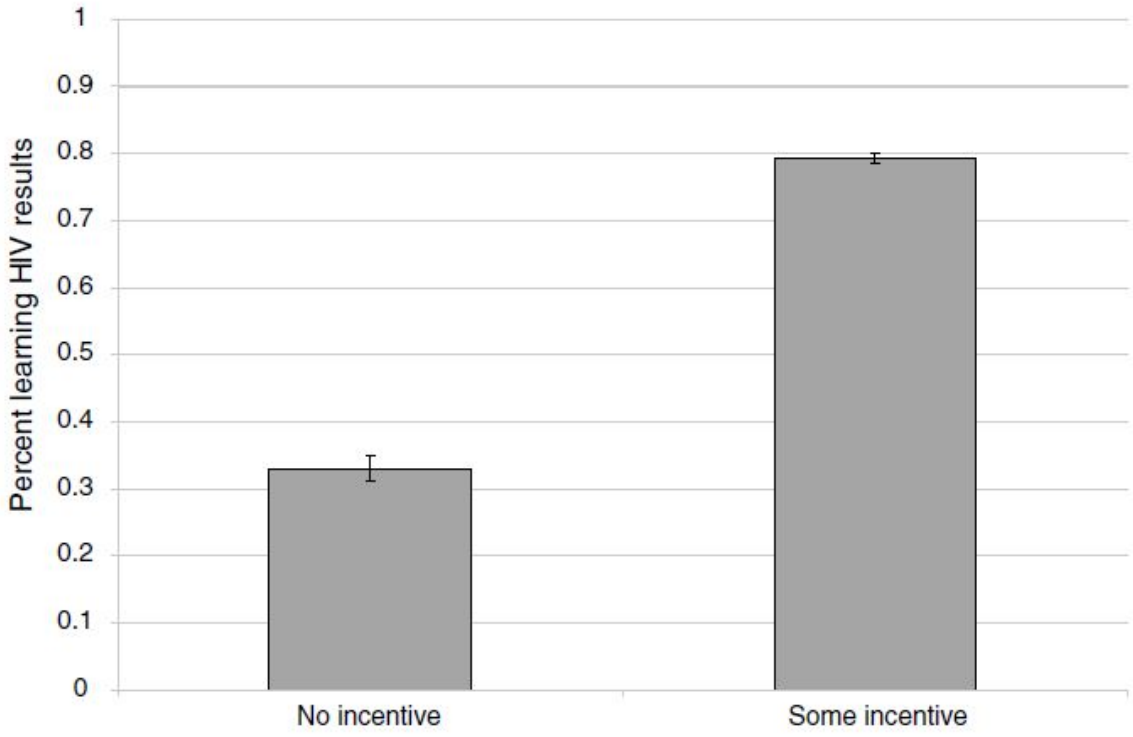

Figure: Visual representation of cash transfers on learning HIV test results.

Results

- Even small incentives were effective
- Any incentive increases learning HIV status by 43% compared to the control (mean 34%)
- Next she looks at the effect that learning HIV status has on risky sexual behavior

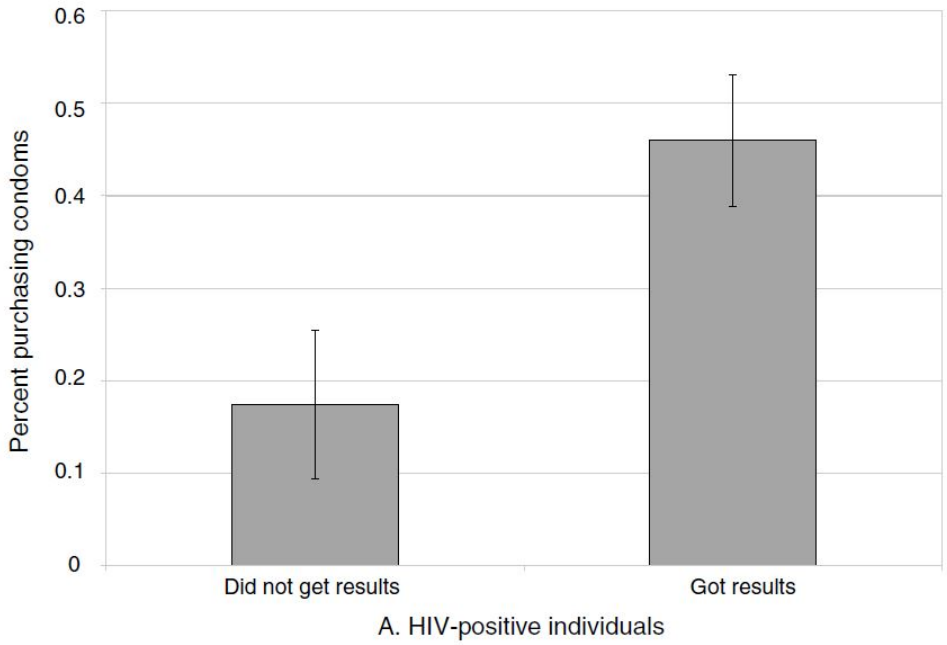

Figure: Visual representation of cash transfers on condom purchases for HIV positive individuals.

Table: Reactions to Learning HIV Results among Sexually Active at Baseline

Dependent variables:	**Bought condoms**		**Number of condoms bought**	
	OLS	**IV**	**OLS**	**IV**
Got results	-0.022	-0.069	-0.193	-0.303
	(0.025)	(0.062)	(0.148)	(0.285)
Got results × HIV	0.418***	0.248	1.778***	1.689**
	(0.143)	(0.169)	(0.564)	(0.784)
HIV	-0.175**	-0.073	-0.873	-0.831
	(0.085)	(0.123)	(0.275)	(0.375)
Controls	Yes	Yes	Yes	Yes
Sample size	1,008	1,008	1,008	1,008
Mean	0.26	0.26	0.95	0.95

Results

- For those who were HIV+ and got their test results, 42% more likely to buy condoms (but shrinks and becomes insignificant at conventional levels with IV).
- Number of condoms bought – very small. HIV+ respondents who learned their status bought 2 more condoms