Randomization

Slides for AERC Conference

Jeremy Magruder

UC-Berkeley

May 2009

A Common Econometric Problem

- Much of development is, ultimately, about what programs work and what programs don't
- That is, for any given program, we want to know the true counterfactual: that is, how people who receive the program would have looked if it weren't for the program
- In principle, this is impossible to observe:
  - With any program, people either receive the program or they don't
  - We observe outcomes for program recipients and outcomes for non-recipients
  - It is impossible, at any point in time, to observe the true counterfactual

A Concrete Example

- A common problem is that people don't adopt technologies that would be profitable for them
- These may include hybrid varieties, fertilizers, capital, or export-oriented crops
- Lets think about export-oriented crops:
- Many reasons for underadoption
    - lack of knowledge
    - lack of connections to international market
    - up-front costs/capital constraints
    - etc.

Extension groups

- Extension agencies which teach about crops, provide connections to international markets, potentially offer loans may help with all of these
- Ashraf, Gine, and Karlan (2008) want to learn whether a particular extension agency has helped Kenyan farmers adopt export crops
- These include French Green Beans, Baby Corn, and Passion Fruit
- Much more profitable in some areas than traditional crops
  - Also want to know whether adopting farmers are better off
- This agency (DrumNet) has several offices throughout Kenya, works with farmer self-help groups (SHGs)
  - DrumNet offers both extension and loan services, and intermediates between farmers and exporters

A simple model of extension

- Suppose Y_i is the fraction of fields that farmer i uses on export crops.
- Each farmer might have a different willingness to use export crops, μ_i, which happens no matter what.
- Then, if there were no DrumNet, $Y_i = \mu_i$. However, DrumNet increases farmer willingness to grow export crops. For farmers with access to DrumNet, $Y_i = \beta^{DN} + \mu_i$
- While for those without access, $Y_i = \mu_i$

Econometric Problem Applied Here

- What we can observe in the data is which farmer SHGs are working with DrumNet and which ones are not
- Could look at crop choice for DrumNet SHGs versus non-DrumNet SHGs.

$$\Delta_1 = \bar{Y}^{DN} - \bar{Y}^{NDN}$$

- However, DrumNet seeks out interested SHGs, and interested SHGs may seek out DrumNet
  - These SHGs may grow more Export Crops regardless of DrumNet's assistance. If N_{DN} farmers receive DrumNet, then

$$\begin{aligned}
\bar{Y}^{DN} &= \beta^{DN} + \frac{1}{N_{DN}} \sum \mu_i = \beta^{DN} + \mu_1 \\
\bar{Y}^{NDN} &= \frac{1}{N - N_{DN}} \sum \mu_i = \mu_2 \\
\Delta_1 &= \beta^{DN} + \mu_1 - \mu_2
\end{aligned}$$

Econometric problem (2)

- This means that

$$E\left[\bar{Y}|DN=1\right]=\beta_{DN}+E\left[\mu|DN=1\right]=\beta_{DN}+\mu_1$$

- so that when we regress

$$Y_i=\bar{\mu}+\beta*DN_i+\mu_i$$

- μ_i is correlated with Drum Net status and we misestimate our results
- We could try including a lot of control variables... but, μ_1 is fundamentally unobservable. A lot of things which are probably important, like a farmer's ingenuity and openness to new ideas, are probably not that strongly correlated with many observables and would be left in the error.

Randomization

- If we could get DrumNet to randomly offer their services to some self-help groups and not offer them to others, we can avoid this problem.
- That is, the whole econometric problem was created because there was non-random selection into treatment
- If selection is random, then

$$E\left[\bar{Y}|DN=1\right]=\beta_{DN}+E\left[\mu|DN=1\right]=\beta_{DN}+\bar{\mu}$$

- So that we get consistent estimates by regressing

$$Y_i=\bar{\mu}+\beta*DN+\mu_i$$

- AGK follow this approach. They have 36 farmer SHGs groups, which they divide into three groups:
  - One Control group (C), who gets no treatment
  - A Full Treatment Group (T), which gets all of Drum Nets services (both extension and a loan)
  - and an Extension treatment group (E), which gets only extension services.

Randomization Comments

- With these three groups, they should be able to proceed simply.
- If these groups are chosen truly at random then

$$
\begin{aligned}
E[\bar{Y}|E] &= \beta^E + E[\mu|E] = \beta^E + \bar{\mu} \\
E[\bar{Y}|T] &= \beta^T + E[\mu|T] = \beta^T + \bar{\mu} \\
E[\bar{Y}|C] &= E[\mu|C] = \bar{\mu}
\end{aligned}
$$

- So that the effect of extension services, $\beta^E = \bar{Y}^E - \bar{Y}^C$
- and the effect of all of DrumNet's services, $\beta^T = \bar{Y}^T - \bar{Y}^C$

Randomization Comments

- It's worth recalling the questions that led us to this point.
- We wanted to know why it is that people under-adopt new and profitable technologies. Can we learn that from this?
  - We have two treatments, basically. We can learn whether DrumNet's extension services lead to more adoption
  - We can learn whether DrumNet's Extension services + loan services lead to more adoption
  - We can learn whether DrumNet's Extension services + loan services have a larger effect than just the extension services
- But, we don't learn which extension services matter
  - Don't learn whether it's an issue of access to international markets, or knowledge about new crops, etc.
  - These might be really important for policy, or for learning how this expt would play out in a different context
- Also don't learn whether loans – by themselves – would be just as effective
  - Only observe loans for people who received extension services

Same Comments, Algebraically

- To be careful, suppose a the effect of extension services, $\beta^E = \beta_A + \beta_K$
- suppose that the effect of full treatment, $\beta^T = \beta_A + \beta_K + \beta_L + \beta_L(\beta_A + \beta_K)$.
- Then if we take $\bar{Y}^T - \bar{Y}^E$, we are left with

$$\begin{aligned} \beta^T - \beta^E &= \beta_A + \beta_K + \beta_L + \beta_L(\beta_A + \beta_K) - \beta_A + \beta_K \\ &= \beta_L + \beta_L(\beta_A + \beta_K) \end{aligned}$$

Sample Size Constraints

- Obviously, with more treatment groups, we could learn more about the mechanisms which are important here
- Here, we have about 400 farmers, but only 36 SHGs.
  - Not feasible to provide different treatments to different farmers in the SHG
  - loans are group loans, extension meetings joint, etc.
- This means that the "treatment" happens to each group, not to individuals – outcomes are correlated
- Very hard to put fewer groups in each treatment

Sample Size Constraints (2)

- Two practical considerations:
  - Since observations within a group are correlated (treatment is correlated), have to cluster standard errors at SHG level
  - Important to keep sufficiently many clusters – theory on clustered standard errors all assumes a large number
  - In this case, may not be enough for reliable estimated errors
- With small sample sizes, increased probability that treatment group is different from control group
  - more likely that treatment/control could be manipulated
  - more likely, also, that just by chance treatment or control are different
  - Can't test whether treatment and control are different on unobservable dimensions
  - can test observable dimensions

		Means			
	N	All	Control	Treatment	p-value
Current Number of Members	36	28.7	31.4	27.3	0.51
		(17.5)	(19.6)	(16.6)	
Age of SHG (months)	36	4.77	4.99	4.66	0.85
		(4.9)	(3.9)	(5.4)	
SHG has social activities	36	0.53	0.75	0.42	0.06*
		(0.5)	(0.5)	(0.5)	
Fee Contributions to SHG	36	103	87.5	111	0.55
		(106.0)	(56.9)	(124.0)	
SHG has a bank account	36	0.64	0.67	0.63	0.81
		(0.5)	(0.5)	(0.5)	
Main Road Paved	36	0.86	1	0.79	0.09*
		(0.4)	(0.0)	(0.4)	
KM to main market	36	5.82	5.08	6.19	0.39
		(3.6)	(3.2)	(3.8)	
Time to the main market	36	41.5	22.5	51	0.09*
		(47.1)	(16.0)	(54.6)	

Differences in Observable Dimensions

- It's true that, on average, randomly selected towns should be the same.
- However, the probability of them being different from each other in sample is pretty high when we have only a few treatment and control group members (like in this case)
- Even if our observables were balanced, we still might worry that there was an important unobservable difference just due to chance
- Randomization only creates identical groups in expectation – this is a large N statement
- This means randomization, alone, may not be enough
- In particular, although $\bar{Y}^T - \bar{Y}^C$ on average gives us a consistent estimate of the program, it may be biased if treatment and control are different due to chance

Difference-in-Differences

- If we are lucky enough to have data from before DrumNet started it's services, we can try one approach to correct this.
- Suppose the difference between DrumNet SHGs and non-DrumNet SHGs is something that stays the same over time (that is, suppose $E[\mu|T] = \mu_T, E[\mu|C = \mu_C], \mu_T \neq \mu_C$)
- That is, it may be that DrumNet SHGs always grow more green beans, and that this happens the same both before and after DrumNet exists.
- We could try
$$\Delta_2 = \bar{Y}^T_{Post} - \bar{Y}^C_{\Pr e}$$
- But, if something is changing over time – for example, crop prices change, or local awareness of export crops, then that would give us
$$\beta^T + \mu_T + \alpha^{Post} - \mu_T - \alpha^{\Pr e}$$
- And we would be left with
$$\beta^T + \alpha^{Post} - \alpha^{\Pr e}$$

Difference-in-Differences

- A solution is to do differences in differences:

$$DD = \left(\bar{Y}^{T}_{Post} - \bar{Y}^{T}_{Pre}\right) - \left(\bar{Y}^{C}_{Post} - \bar{Y}^{C}_{Pre}\right)$$

$$\begin{aligned} DD &= \left(\beta^{T} + \mu_{T} + \alpha^{Post} - \mu_{T} - \alpha^{Pre}\right) - \left(\mu_{C} + \alpha^{Post} - \mu_{C} - \alpha^{Pre}\right) \\ &= \beta^{T} \end{aligned}$$

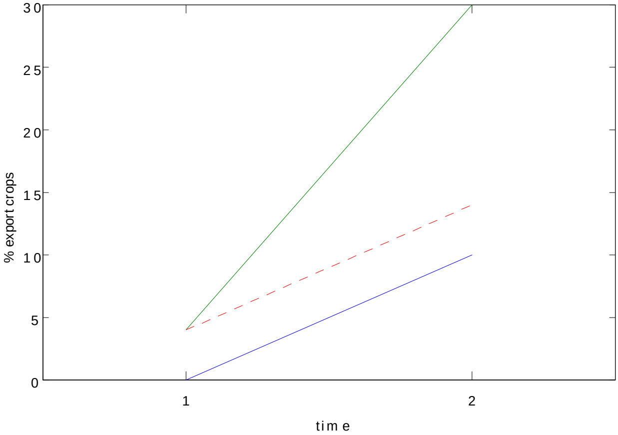
- We can also write this as a regression

$$Y_i = \mu + \beta_1 T_i + \beta_2 Post_t + \beta^{T} T_i * Post_t + \varepsilon_{it}$$

Difference-in-Differences (2)

- The difference-in-difference approach assumes that:
  - Some things may change over time (i.e. knowledge/availability of export crops)
  - Some things may be different between treatment and controls (i.e. distance to market)
  - But, the things which change over time are not different between treatment and control groups
  - idea is that, were it not for treatment, the treatment group would have experienced the same change in take-up as the control group
- This approach is particularly convincing when we've randomized treatment and control – now, the only differences between these guys are because we got unlucky. We can hope that the ways we've gotten unlucky are constant over time
  - If we didn't randomize, would worry that treatment farmers were more eager to adopt tech, then if tech became more available they would systematically adopt, control would not

Difference-in-Differences (3)



Difference-in-Differences (4)

	Any Export Crop	% land Export	Log HH income
Post	-0.004	-0.079***	-0.109
	(0.06)	(0.02)	(0.10)
Post*treatment	0.192***	0.043*	0.087
	(0.07)	(0.02)	(0.11)
N	1983	1779	1566
R-square	0.27	0.13	0.16

Difference-in-Differences (4)

	Any Export Crop	% land Export	Log HH income
Post	-0.004	-0.079***	-0.11
	(0.06)	(0.02)	(0.10)
Post*Credit	0.226***	0.049*	0.011
	(0.08)	(0.03)	(0.12)
Post*No Credit	0.159**	0.037	0.162
	(0.07)	(0.03)	(0.12)
N	1983	1779	1566
R-square	0.29	0.7	0.12

What are these numbers telling us?

- We asked how groups that were randomly assigned treatment performed relative to groups which were not
- However, farmers within those groups could have chosen to join DrumNet or not
- Some farmers may be disinterested in export crops; others may already be growing and marketing them on their own
- In fact, only about 1/3 of eligible farmers in treatment group joined DrumNet
- That means that our estimated effect is the average effect for some farmers who chose to receive DrumNet, and some who did not.
- This is $E\left[\bar{Y}_s^T - \bar{Y}_s^C\right]$, the "Intention to Treat" effect

Average Treatment Effects

- Remember, we wanted to know how much people would gain, on average, from Extension services
- This is not what we just measured.
- Suppose each farmer i would experience some gain from DrumNet
- Then, under DrumNet, farmer i would produce Y_i^{DN}; without DrumNet he would produce Y_i^{NDN}
- We may want to know, on average, how much better off would people be if they all used DrumNet?
- This is $E\left[Y_i^{DN} - Y_i^{NDN}\right]$, the average treatment effect
- That is not observed, because some people choose to receive Drum Net and some do not.

Heterogeneity in treatment effects

- We observe treatment outcomes only for the treated.
- Income did not increase on average in treated communities...
  - but it may have increased for treated individuals
  - Or, it may have increased for some group of treated individuals
- One possibility: use eligibility of treatment as an instrument for actually being treated
  - Assuming farmers who don't join DrumNet are unaffected, this tells us the total effect on farmers who join
  - $E\left[Y_i^{DN} - Y_i^{NDN} | i \text{ joins DrumNet}\right]$
  - Still not the average treatment effect - Average Treatment on The Treated
  - Amounts to scaling the coefficients above by the participation rate
  - In this case, would suggest that 60% of DrumNet participants started growing export crops

Heterogeneity in treatment effects

- Can also examine whether some groups who might particularly benefit did
- We don't know who would join in Control groups
- But, we do know which farmers in Treatment and Control groups were already growing Export Crops before DrumNet
- Farmers who hadn't yet adopted Export Crops may benefit particularly

	% Land Export		Log HH income	
Post	-0.099***	-0.056	-0.129	-0.132
	(0.02)	0.033	(0.09)	(0.18)
Post*Treatment	-0.02	0.09**	-0.32	0.319*
	(0.03)	0.04	(0.12)	(0.18)
Export at Baseline?	Yes	No	Yes	No
N	818	909	764	744
R-square	0.18	0.14	0.2	0.19

Conclusions

- Difficult to identify effect of programs as true counterfactual not observed
- Randomization can help a lot with this
- With Small Samples, Diff-in-Diffs combined with randomization may be more satisfying
- Heterogeneity in treatment effects: important to keep track of what our estimates are saying, precisely

Study Epilogue/Ethics

- In this study, DrumNet was ultimately unable to buy the product after 1 year as Europe changed import reqs
- Most tech adopters switched back to traditional crops after that year
- Many took a loss from the program in the end
- This highlights an important point: any intervention can induce risks
    - In this case, those risks turned out badly
    - Important for researchers investigating any intervention to consider possible risks – in this case, probably nothing could be done

Ethics of Randomization

- Though almost any intervention contains risks, randomization has some particular ethical concerns
- Some people are denied treatment, while others are offered it
- If the program is really helpful and we can afford to give it to the entire study group, it may be difficult to justify not giving it to all
  - However, we (realistically) almost always have limited funds as researchers.
  - Low ethical costs to monitoring those who we cannot afford to treat anyway
  - Randomization may even be seen as a "fair" way to determine who receives program and who does not
- Even if we can afford to treat full sample, we often consider rolling out treatment in random waves, if cost of delaying treatment is not too high

When can we randomize?

- Natural Experiments with Purposeful Randomization
  - Governments sometimes implement policies in a random way
  - e.g. Political appointment reservations in India, early Progresa beneficiaries
- Natural Experiments created by human error
  - Brazilians with a Spanish name dropped from a program b/c their names included a letter not in Portuguese
- NGOs/Governments who want to learn the effectiveness of a program
  - e.g. DrumNet, class size experiments, teacher effectiveness experiments, etc.
  - Most randomized studies belong here