

Bài giảng 7: Ôn tập

Lê Việt Phú

9/7/2018

So sánh nhóm hưởng lợi và nhóm đối chứng

- ▶ Comparing outcomes for the treated and untreated often yields incorrect estimates

$$E[Y|D = 1] - E[Y|D = 0]$$

$$E[Y_1|D = 1] - E[Y_0|D = 0]$$

$$E[Y_1|D = 1] - E[Y_0|D = 0] + (E[Y_0|D = 1] - E[Y_0|D = 1])$$

which rearranges to:

$$E[Y_1|D = 1] - E[Y_0|D = 1] + E[Y_0|D = 1] - E[Y_0|D = 0]$$

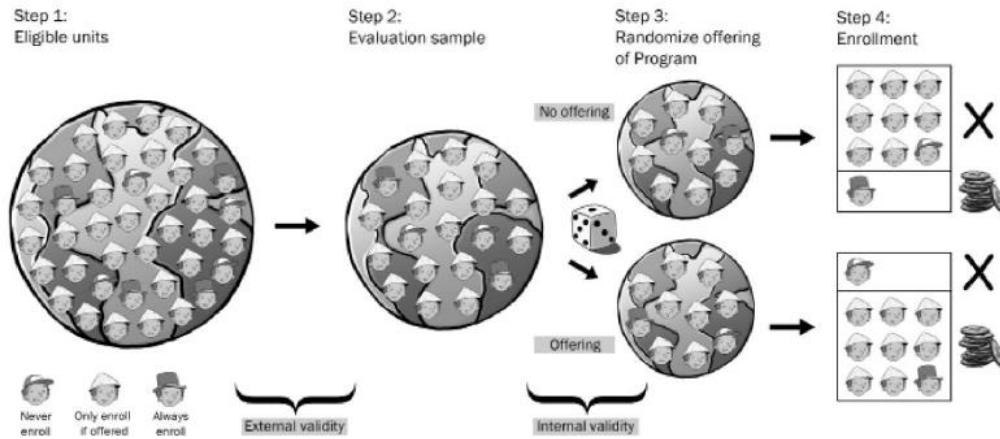
which consists of the Average Treatment Effect on the Treated (ATET) and Selection Bias:

$$\underbrace{E[Y_1 - Y_0|D = 1]}_{\text{ATET}}$$

$$+ \underbrace{E[Y_0|D = 1] - E[Y_0|D = 0]}_{\text{Selection Bias}}$$

Phân biệt các khái niệm ATE, ITE, TOT

Ngẫu nhiên hóa đề xuất tham gia



	Group offered treatment	Group not offered treatment	Impact
	% enrolled = 90% Average Y for those offered treatment = 110	% enrolled = 10% Average Y for those not offered treatment = 70	$\Delta\%$ enrolled = 80% $\Delta Y = ITT = 40$ ToT = $40/80\% = 50$
Never enroll			—
Only enroll if offered the program			
Always enroll			—

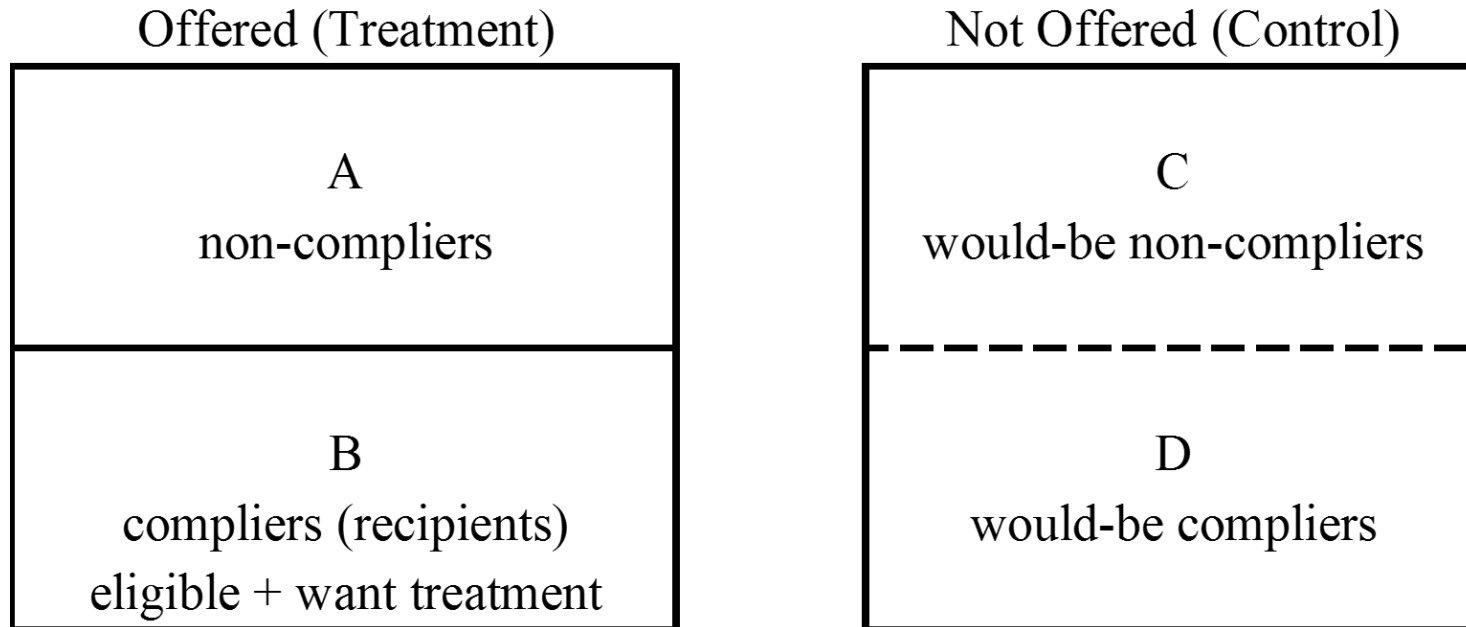
ITT: Intention to Treat.
Impact of program on those who are offered the treatment, *regardless of whether they actually enroll*

TOT: Treatment on Treated.
Impact of program on those who are offered treatment & *who actually enroll*

Source: Gertler 2011

- ITT: Tác động của chương trình đối với nhóm được đề xuất tham gia, bất chấp việc trên thực tế họ có tham gia hay không
- TOT: Tác động của chương trình đối với nhóm được đề xuất tham gia và trên thực tế có tham gia.

So sánh ITE và TET



- ITE so sánh A & B với C & D.
- So sánh B với C & D có hàm chứa tác động chệch mang tính lựa chọn ngay (selection bias) cả khi việc lựa chọn tham gia chương trình được thực hiện ngẫu nhiên
- TET so sánh B với D, tuy nhiên làm thế nào để đảm bảo tuân thủ trong nhóm đối chứng?

Tác động muốn đánh giá sẽ quyết định thiết kế nghiên cứu

- Tác động can thiệp trung bình (Average Treatment Effect): có thể trực tiếp sử dụng thử nghiệm ngẫu nhiên.
- Tác động lên nhóm được đề xuất tham gia (Intention to Treat Effect):
 - Lấy mẫu ngẫu nhiên cho cả quần thể.
 - Tiến hành lựa chọn mẫu trong quần thể các nhóm đã được lựa một cách ngẫu nhiên
 - ITE được tính bằng sự khác biệt giữa kết quả của nhóm được đề xuất tham gia so với nhóm không.
- Tác động can thiệp lên đối tượng tham gia (Treatment Effect on Treated):
 - Chỉ sử dụng một nhóm các cá nhân được đề xuất tham gia và sẽ làm theo đúng yêu cầu. Đó là nhóm những người tuân thủ.
 - Phân bổ ngẫu nhiên việc tham gia chương trình trong nhóm tuân thủ này.
 - TET được tính bằng sự khác biệt về kết quả giữa nhóm có tham gia và không tham gia của những người tuân thủ.
 - Không dễ ước lượng TET trên thực tế vì yêu cầu phải xác định được nhóm những người tuân thủ và yêu cầu họ đăng ký vào chương trình nghiên cứu thử nghiệm. Việc thực hiện tương tự như bốc thăm.

Sức mạnh thống kê, cỡ mẫu và tác động tối thiểu có thể phát hiện được

Sức mạnh kiểm định thống kê

- Độ mạnh của một kiểm định là khả năng (xác suất) bác bỏ một giả thuyết thông kê (giả thuyết không – Null) khi giả thuyết thay thế là đúng.
- Khi nghiên cứu so sánh 2 nhóm, độ mạnh của kiểm định là xác suất bác bỏ giả thuyết không là hai nhóm có cùng một giá trị trung bình (bằng trung bình của quần thể), do đó kết luận là không có sự khác biệt giữa giá trị trung bình của hai quần thể, khi mà trên thực tế có sự khác biệt ở một mức độ nhất định.
- Do đó nó là xác suất kết luận đúng là hai nhóm khác biệt nhau.

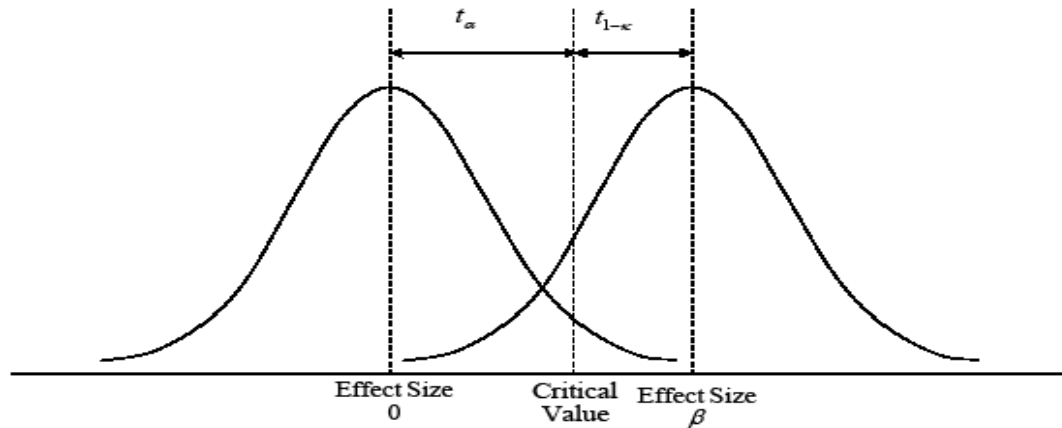
Sức mạnh kiểm định thống kê

	Test result	
	“Reject Null,” Find an effect!	“Fail to Reject Null” Conclude no effect.
Truth: There is an effect	Great! (κ)	“Type II Error” (low power $1 - \kappa$)
Truth: There is NO effect	“Type I Error” (test size α)	Great! ($1 - \alpha$)

- Probability of Type I error is “size” of test, α (typically 0.05)
- Probability of Type II error is $(1 - \kappa)$, where κ is “power” of the test

Sức mạnh thống kê và mức ý nghĩa

Figure 1



Đồ thị trái là phân phối của beta mũ theo giả thuyết không là nó bằng 0.

Đồ thị phải là phân phối của beta mũ nếu độ lớn thực tế là beta.

Mức ý nghĩa được nhận diện là vùng đuôi bên phải của phân phối bên trái.

Sức mạnh thống kê là vùng đuôi bên trái của phân phối bên phải.

(source: Duflo & Kremer 'Toolkit')

Cỡ mẫu tối thiểu

$$N > \frac{\sigma^2}{\left(\frac{ETE}{\left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)}} \right)^2}$$

Bạn có thể chấp nhận một cỡ mẫu nhỏ nếu:

- Kỳ vọng tác động can thiệp là lớn
- Kết quả có độ dao động nhỏ
- Nhóm tham gia và đối chứng là các nhóm tương đồng về số quan sát ($p=.5$)
- Sẵn lòng chấp nhận mức ý nghĩa thấp và độ mạnh thống kê thấp.

Thiết kế can thiệp theo nhóm (cluster design)

Nhận xét về sự khác biệt giữa “tác động tối thiểu có thể phát hiện được” – tác động thực nhỏ nhất *mà một thử nghiệm có thể phát hiện được với xác suất cao.*

Không có thiết kế theo nhóm:

$$MDE > \left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)}} \sqrt{\frac{\sigma^2}{N}}$$

Với thiết kế theo nhóm:

$$MDE > \left(t_{1-\kappa} + t_{\alpha/2} \right) \sqrt{\frac{1}{p(1-p)J}} \sigma \sqrt{\rho + \frac{1-\rho}{n}}$$

(J là số nhóm cùng kích cỡ, ρ là tương quan nội nhóm, và n là số quan sát trong một nhóm.)

Thử nghiệm tự nhiên và các phương pháp ước lượng tác động sử dụng dữ liệu quan sát được

Thử nghiệm tự nhiên là gì?

- Nhà nghiên cứu lợi dụng các tình huống can thiệp tự nhiên để lựa chọn nhóm tham gia và nhóm hưởng lợi thay vì quá trình thử nghiệm ngẫu nhiên hóa.
- Nhà nghiên cứu không can thiệp thay đổi được mẫu nghiên cứu, ví dụ phân bổ ai vào nhóm nào, hưởng lợi hay đối chứng.
- Nhà nghiên cứu phải chứng minh rằng quá trình can thiệp tự nhiên tạo ra nhóm hưởng lợi và đối chứng hoàn toàn ngẫu nhiên - tương tự như phân bổ ngẫu nhiên trong một thử nghiệm có kiểm soát.

Tại sao lại sử dụng thử nghiệm tự nhiên?

- Giúp nhà nghiên cứu xử lý 3 vấn đề thực nghiệm:
 1. Vấn đề nhân quả ngược/nội sinh (Reverse Causality/Endogeneity): Một cú sốc ngoại sinh giúp phá vỡ vòng luẩn quẩn nguyên nhân – kết quả.
 2. Khác biệt về các đặc tính không quan sát được: Unobserved Heterogeneity (OV Bias)
 3. Tác động lên các biến giải thích: Thử nghiệm tự nhiên giúp xác lập nhóm hưởng lợi và đối chứng một cách rõ ràng.

Những giả định “có thể coi là ngẫu nhiên”

Kém hiệu lực nhất

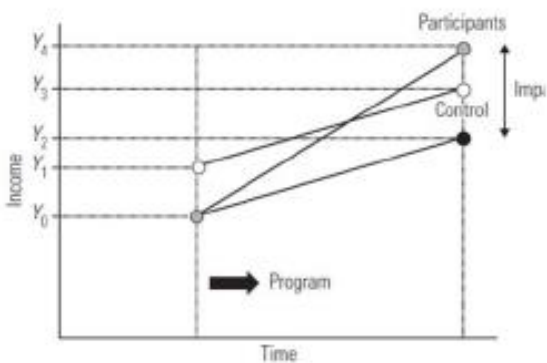
Hiệu lực nhất



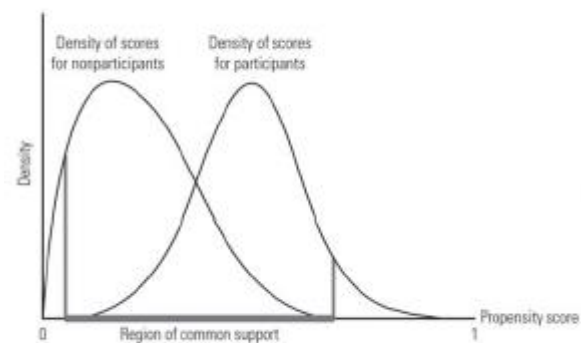
- Giả định “có thể được coi là ngẫu nhiên” càng kém hiệu lực thì nhà nghiên cứu càng phải chắc chắn về mô hình và các nhân tố gây nhiễu (confounding factors) khi đánh giá tác động.
- Do đó, nghiên cứu sử dụng thử nghiệm tự nhiên cần có nền tảng lý thuyết rất vững vàng.

Các phương pháp tiếp cận đánh giá tác động

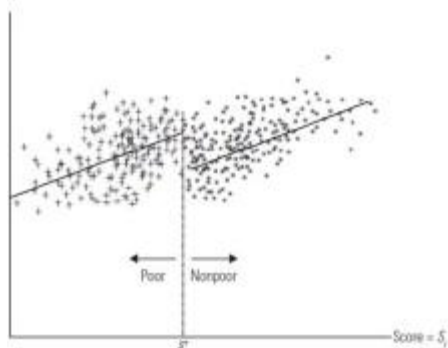
Differences in Differences



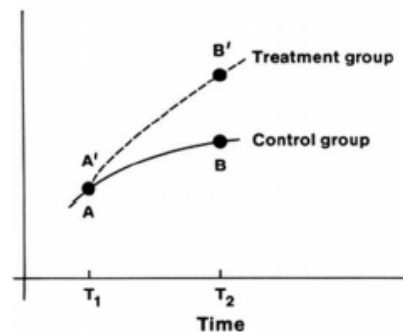
Matching



Regression Discontinuity



Randomized Evaluation



Phương pháp Diff-in-Diff (DD)

- Sử dụng với dữ liệu quan sát được (observational data), không phải dữ liệu thử nghiệm ngẫu nhiên.
- Có thể xử lý được hai vấn đề liên quan đến “phản thực không hợp lệ” (counterfeit counterfactuals)
- DD giả định tồn tại các khác biệt không quan sát được giữa các nhóm, nhưng chúng không thay đổi theo thời gian
 - Khác biệt về giá trị trung bình của kết quả theo nhóm sẽ tự động bị khử đi
- Có thể kết hợp với các phương pháp khác như phương pháp ghép cặp (matching)

Phản thực không hợp lệ là gì?

- Nhóm đối chiếu không hợp lệ sẽ làm cho kết quả bị chệch
- 2 phương pháp có thể dẫn đến phản thực không hợp lệ:
 1. So sánh kết quả của cùng một nhóm tham gia chính sách trước và sau khi thực hiện chính sách
 2. So sánh trực tiếp kết quả của nhóm tham gia và nhóm không tham gia
- Nếu nhóm đối chiếu không hợp lệ, tác động của chính sách bị nhiễu với tác động của các nhân tố khác

Cơ chế của phương pháp DD

- So sánh thay đổi trong kết quả theo thời gian giữa nhóm hưởng lợi và nhóm đối chứng
- Giả định khác biệt không quan sát được nhưng không thay đổi theo thời gian

	<u>Treatment</u>	<u>Control</u>
Before	T_B	C_B
After	T_A	C_A

$$DD = (T_A - T_B) - (C_A - C_B)$$

- *Ước tính tác động như sau:*
 - (1) Tính khác biệt về kết quả của nhóm hưởng lợi
 - (2) Tính khác biệt về kết quả của nhóm đối chứng
 - (3) Tác động của chính sách là khác biệt giá trị (1) so với (2)

Cơ chế của phương pháp DD

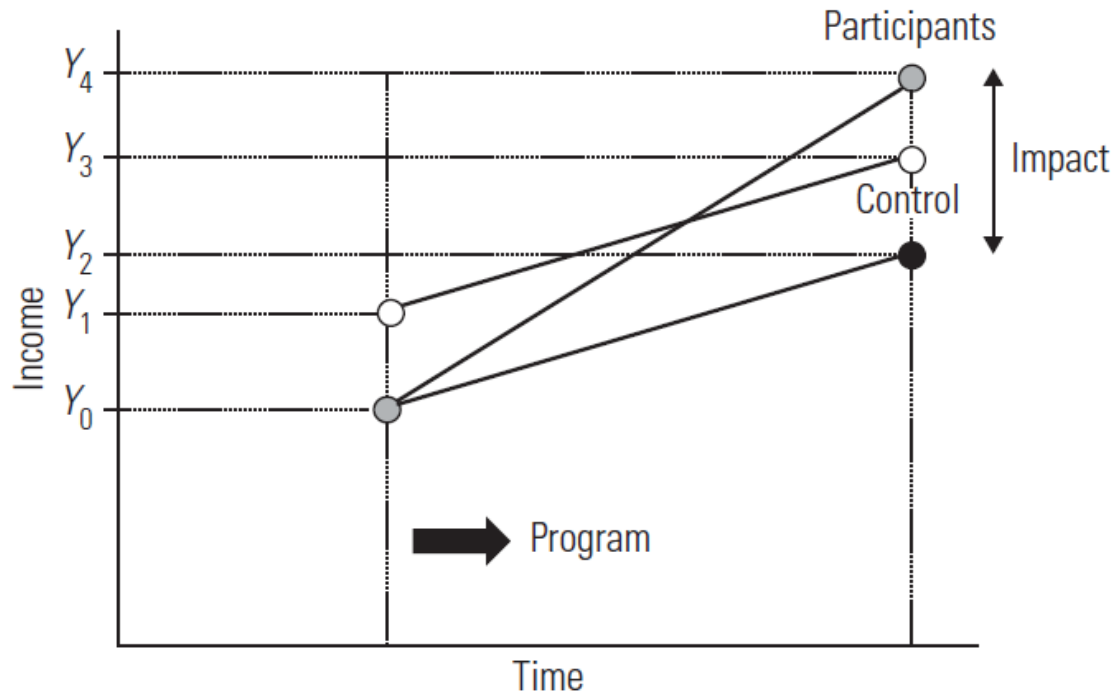
	<u>Treatment</u>	<u>Control</u>
Before	T_B	C_B
After	T_A	C_A

$$DD = (T_A - T_B) - (C_A - C_B)$$

- Kết hợp 2 phản thực không hợp lệ
 - So sánh trước và sau khi thực hiện chương trình của nhóm tham gia
 - So sánh nhóm tham gia với nhóm không tham gia
- **1st difference**: Δ về kết quả trước và sau chương trình của nhóm tham gia
 - Cho phép kiểm soát các vấn đề không thay đổi theo thời gian
- **2nd difference**: Δ về kết quả trước và sau chương trình của nhóm không tham gia
 - Kiểm soát các thay đổi đối với nhóm tham gia khi không có chính sách
 - Để ước lượng phản thực yêu cầu phải có nhóm đối chứng hợp hệ
- Thực hiện DD cho phép kiểm soát những đặc tính hoặc quan sát được hoặc không, và không thay đổi theo thời gian

Giả định song song

- Nếu không có chương trình thì xu hướng thay đổi kết quả của nhóm tham gia và nhóm đối chứng là như nhau
- Để ước lượng phản thực yêu cầu không có các khác biệt thay đổi theo thời gian giữa các nhóm tham gia và đối chứng. Nếu xu hướng giữa các nhóm khác nhau thì ước lượng sẽ bị chệch
- Nếu thu nhập của nhóm đối chứng tăng nhanh hơn thì ước lượng bị chệch xuống
- Giả định song song không thể chứng minh một cách chắc chắn, tuy nhiên có thể sử dụng một số kiểm định để kiểm chứng sự phù hợp



Sử dụng hồi quy để ước lượng

$$Y = \beta_0 + \beta_1 TRT + \beta_2 POST + \beta_3 TRT \cdot POST + \varepsilon$$

	<u>New Jersey</u>	<u>Pennsylvania</u>
Before Increase	$\beta_0 + \beta_1$	β_0
After Increase	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_2$

$$DD = (T_A - T_B) - (C_A - C_B) = [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_1)] - [(\beta_0 + \beta_2) - (\beta_0)] = \beta_3$$

Có thể đưa thêm các biến kiểm soát khác vào mô hình:

$$Y = \beta_0 + \beta_1 TRT + \beta_2 POST + \beta_3 TRT \cdot POST + \beta_4 X + \varepsilon$$