

Mô hình Biến Công cụ (Instrumental Variables Design)

Lê Việt Phú
Chương trình Giảng dạy Kinh tế Fulbright

Ngày 21 tháng 5 năm 2016

Ôn tập mô hình hồi quy đánh giá tác động chính sách

Giả sử chúng ta muốn ước lượng tác động của việc tham gia chính sách tín dụng hộ gia đình lên thu nhập của hộ bằng một mô hình đơn giản sau:

$$Y_i = \alpha X_i + \beta * T_i + \varepsilon_i \quad (1)$$

trong đó:

- ▶ Y_i là thu nhập của hộ.
- ▶ X_i là các đặc tính nhân khẩu học. Ví dụ như tuổi, giới tính, trình độ học vấn của chủ hộ.
- ▶ T là biến chính sách, nhận giá trị 0 và 1 đối với gia đình không và có tham gia.

$$T = \begin{cases} 1 & \text{Tham gia} \\ 0 & \text{Không tham gia} \end{cases}$$

- ▶ ε_i là phần dư.
- ▶ Tham số cần ước lượng β là tác động của việc tham gia chính sách đến thu nhập của hộ gia đình.

Ôn tập lý thuyết hồi quy tuyến tính cổ điển - CLRM

- ▶ Các điều kiện để ước lượng của β trong mô hình (1) là BLUE.
 - ▶ $E[\varepsilon_i] = 0$
 - ▶ $E[\varepsilon_i * T_i] = 0$
 - ▶ $V[\varepsilon_i] = \sigma^2$
 - ▶ $Cov[\varepsilon_i, \varepsilon_j] = 0$

- ▶ Các đặc tính của ước lượng BLUE:
 - ▶ Giá trị tới hạn ước lượng được của β bằng với giá trị chuẩn.
 - ▶ Sai số ước lượng được của β là nhỏ nhất.
 - ▶ Xác suất ước lượng β hội tụ về giá trị chuẩn.

Các mô hình đánh giá tác động chính sách đã học

- ▶ Ngẫu nhiên hoá việc chọn tham gia chương trình (randomization) để đảm bảo nhóm tham gia và nhóm đối chứng hoàn toàn tương đồng và không có hiện tượng tự lựa chọn mẫu \Rightarrow Tác động can thiệp của chính sách là khác biệt về thu nhập giữa hai nhóm.
- ▶ Sử dụng tình huống bán thực nghiệm hay thực nghiệm tự nhiên (quasi/natural experiment): Không đảm bảo nhóm tham gia và nhóm đối chứng hoàn toàn tương đồng, nhưng đảm bảo không có vấn đề tự lựa chọn mẫu. Khi này sử dụng PSM hoặc DiD có thể tìm được nhóm đối chứng và tính được tác động can thiệp của chính sách.

Tại sao khó đảm bảo việc tham gia chính sách là hoàn toàn ngẫu nhiên?

- ▶ Chính sách luôn có mục tiêu cụ thể, ví dụ hướng vào đối tượng ưu tiên thay vì cho toàn bộ dân số (purposive placement).
- ▶ Tự lựa chọn mẫu (self selection): những hộ thực sự cần thiết tham gia chưa chắc đã là những hộ được tham gia chính sách, hoặc ngược lại, do những nguyên nhân không quan sát được.
- ▶ **Hiện tượng tham gia chính sách không ngẫu nhiên còn được gọi là vấn đề biến nội sinh (endogeneous variables). Đây là vấn đề đặc biệt quan trọng trong các mô hình định lượng.**

Hậu quả của việc ước lượng mô hình (1) nếu việc tham gia chính sách là không ngẫu nhiên

$$\text{plim } \hat{\beta} = \beta + \frac{\text{Cov}(\varepsilon * T)}{\text{Var}(T)}$$

- ▶ Nếu biến chính sách không ngẫu nhiên $\Rightarrow \text{cov}(T, \varepsilon) \neq 0$, và ước lượng của β sẽ bị chệch và không nhất quán.
- ▶ Hướng chệch (lên hay xuống) phụ thuộc vào tương quan giữa phần dư với biến chính sách. Nếu chỉ hộ giàu được tham gia chính sách (ε_i lớn khi $T = 1$) thì ước lượng tác động chính sách sẽ bị chệch lên. Khi này kết luận chính sách có tác động tích cực bị phóng đại so với thực tế.
- ▶ Sử dụng phương pháp biến công cụ (Instrumental Variables) để xử lý vấn đề biến nội sinh.

Một số tình huống nghiên cứu

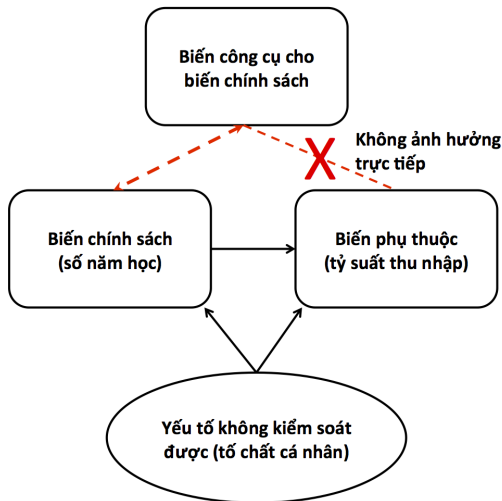
- ▶ Giả sử chúng ta muốn đánh giá tác động của chính sách cho vay tín dụng ưu đãi đến phúc lợi (thu nhập của hộ). Có lý do để cho rằng việc tham gia chính sách là không ngẫu nhiên. Ví dụ gia đình nào có khả năng vay vốn là những hộ có quan hệ tốt với chính quyền, có phương án sử dụng vốn vay hiệu quả, có tài sản thế chấp... Nếu sử dụng hồi quy OLS thì khả năng ước lượng sẽ bị chệch lên do tương quan dương giữa biến chính sách và biến dư (thu nhập).
- ▶ Đánh giá tác động của số năm đi học đến thu nhập. Số năm đi học có khả năng là biến nội sinh do nhiều nguyên nhân: (1) khả năng tài chính và trình độ học vấn của phụ huynh, (2) tố chất cá nhân của người học, (3) thu nhập hay kỳ vọng thu nhập của người học. Các nhân tố trên đều có tương quan dương với thu nhập. Do đó nếu sử dụng hồi quy OLS thì ước lượng sẽ bị chệch lên, dẫn đến kết luận các chính sách giáo dục có hiệu quả cao hơn thực tế.

Vấn đề biến chính sách nội sinh trong các tình huống khác

- ▶ Đánh giá tác động của chính sách bảo hiểm y tế lên thu nhập của nông hộ.
- ▶ Chính sách hỗ trợ vốn cho doanh nghiệp trong giai đoạn khủng hoảng kinh tế.
- ▶ Chương trình cung cấp nước sạch đến cho người dân ảnh hưởng như thế nào đến thu nhập và phúc lợi của hộ sử dụng.
- ▶ Chương trình hỗ trợ đào tạo dạy nghề ảnh hưởng thế nào đến thu nhập của người lao động.

Khung phân tích của phương pháp biến công cụ

Do vấn đề biến nội sinh, không thể ước lượng trực tiếp được tác động của biến chính sách T trong mô hình (1) \Rightarrow Phải ước lượng gián tiếp tác động của tham gia chính sách đến thu nhập. Giả sử tồn tại một biến Z sao cho biến Z ảnh hưởng đến việc lựa chọn tham gia chính sách nhưng không ảnh hưởng đến kết quả của chính sách.



Khung phân tích của phương pháp biến công cụ

Với điều kiện:

- ▶ Z có tương quan với việc tham gia chính sách, $cov(T, Z) \neq 0$.
- ▶ Z không tương quan trực tiếp đến biến kết quả, $cov(Z, \varepsilon) = 0$ (còn gọi là điều kiện loại trừ, exclusion restriction.)

Khi đó có thể dùng biến Z để ước lượng tác động của chính sách bằng hồi quy 2 bước (2-stage least square, 2SLS):

- ▶ Bước 1: Hồi quy $T_i = \gamma Z_i + \phi X_i + u_i$, ước lượng \hat{T}_i
- ▶ Bước 2: Hồi quy $Y_i = \alpha X_i + \beta \hat{T}_i + \varepsilon_i$

$$\beta_{IV} = \frac{cov(Y, Z)}{cov(T, Z)}$$

Khung phân tích của phương pháp biến công cụ

Hiểu thế nào về phương pháp biến công cụ?

- ▶ Bước 1: Tác động đến Y qua 2 hình thức là tác động trực tiếp (cần ước lượng), và gián tiếp (do T và Y cùng đồng tương quan với nhân tố không kiểm soát được). Mục tiêu là ước lượng tác động trực tiếp.

Do Z ngoại sinh với Y nên dùng tương quan giữa T và Z (\hat{T}_i) cũng ngoại sinh với Y . Bản chất của bước 1 là tách phần tác động tác động trực tiếp của chính sách đến Y thông qua Z .

- ▶ Bước 2: Dùng biến \hat{T}_i được tạo từ bước 1 để ước lượng tác động trực tiếp của chính sách lên thu nhập.

Lựa chọn biện công cụ như thế nào?

Biện công cụ phải thoả mãn 2 điều kiện:

- ▶ Tương quan với tình trạng tham gia chính sách.
- ▶ Không tương quan với phần dư của biện phụ thuộc.
- ▶ Rất khó tìm được biện thoả mãn cả hai điều kiện trên. Các biện công cụ thường được sử dụng là các đặc tính địa lý như khoảng cách, hay các thay đổi có yếu tố bất ngờ như các hiện tượng thời tiết cực đoan, thiên tai, hay các chính sách vĩ mô của chính phủ.

Một số ví dụ về biến công cụ

- ▶ Đánh giá tác động của chương trình đào tạo để giúp người thất nghiệp. Việc tham gia chương trình là không ngẫu nhiên. Cần biến công cụ tương quan với việc tham gia, nhưng không trực tiếp tương quan với xác suất xin được việc. Dùng khoảng cách quan sát được giữa nhà với trung tâm đào tạo làm biến công cụ.
- ▶ Nghiên cứu về thu nhập và nội chiến (Miguel et al 2005, JPE). Thu nhập ảnh hưởng đến cạnh tranh tài nguyên và xung đột. Tuy nhiên thu nhập là biến nội sinh. Dùng thay đổi lượng mưa bất thường làm biến công cụ.
- ▶ Nghiên cứu về tác động lâu dài của bom Mỹ đến tăng trưởng kinh tế ở VN (Miguel, JDS). Cường độ ném bom là biến nội sinh, và tăng ở những điểm gần vĩ tuyến 17. Do đó dùng khoảng cách từ các tỉnh đến vĩ tuyến 17 làm biến công cụ.
- ▶ Nghiên cứu về tỷ suất thu nhập của số năm đi học. Sử dụng thời gian sinh theo quý để làm biến công cụ cho biến chính sách là số năm đi học.

Khác biệt giữa IV với hồi quy rút gọn (reduced-form regression)

Tại sao không sử dụng Z_i thay cho biến chính sách T_i và ước lượng phương trình hồi quy (1) như sau:

$$Y_i = \alpha X_i + \beta * Z_i + \varepsilon_i$$

mà phải dùng hồi quy 2SLS?

Thực hành phương pháp biến công cụ

- ▶ STATA file hh_98.dta, IV-2016.do, ivLog.smcl.
- ▶ STATA code: ivregress (biến nội sinh là biến liên tục), treatreg (sử dụng khi biến nội sinh là biến trạng thái tham gia chính sách, 0 1), xtivreg (cho dữ liệu bảng).

Các kiểm định trong phương pháp biến công cụ:

- ▶ Kiểm định Wu-Hausman kiểm định tính nội sinh của biến chính sách T.
- ▶ Biến công cụ yếu (weak instruments): 1st-stage F-stat > 10 (Stock and Yogo, 2005).
- ▶ Điều kiện loại trừ ($Cov(Z, \varepsilon) = 0$, exclusion restriction) không thể kiểm định được, do đó cần giải thích dựa trên kiến thức và bối cảnh của mô hình.

So sánh các phương pháp

	Random	PSM	DiD	IV
Giả định				
Dữ liệu				
Tác động				
Ưu điểm				
Nhược điểm				