

Trường Chính sách Công và Quản lý Fulbright

Học kỳ Xuân 2019
Kinh tế lượng ứng dụng

BÀI TẬP 5

Tác động của bà mẹ hút thuốc lá đến sức khỏe của trẻ sơ sinh*

Ngày Phát: Thứ ba 23/04/2019

Ngày Nộp: Thứ ba 07/05/2019

Giới thiệu

Trẻ sơ sinh thiếu cân (Low Birth Weight - LBW) gặp phải những vấn đề nghiêm trọng về sức khỏe và phát triển sau này, và có thể gây ra tổn thất lớn cho xã hội. Bài nghiên cứu này đánh giá ảnh hưởng của hành vi hút thuốc của bà mẹ khi mang thai, yếu tố nguy cơ hàng đầu của LBW ở Hoa Kỳ, đối với sức khỏe của trẻ sơ sinh.

Trẻ sơ sinh thiếu cân được định nghĩa có cân nặng khi sinh dưới 2500 gram. Những đứa trẻ này sẽ gặp phải những vấn đề sức khỏe nghiêm trọng hoặc chậm phát triển và có thể gây ra chi phí đáng kể cho xã hội. Ví dụ, chi phí dự kiến khi sinh và chăm sóc ban đầu của em bé nặng 1000 gram khi sinh có thể vượt quá 100.000 đô la (năm 2000) và nguy cơ tử vong trong vòng một năm sau sinh là hơn 20%. Ngay cả đối với những em bé nặng 2000-2100 gram và có tỷ lệ tử vong tương đối thấp, nếu trọng lượng khi sinh tăng thêm 1 pound (454 gram) sẽ giảm tới 10.000 đô la chi phí cho các dịch vụ nội trú trong bệnh viện.

Các nghiên cứu đã thiết lập mối tương quan giữa LBW và các bệnh như huyết áp cao (high blood pressure), bại não (cerebral palsy), điếc (deafness), mù (blindness), hen suyễn (asthma) và bệnh phổi (lung disease) ở trẻ em, cũng như IQ, khả năng học hành, vấn đề hành vi và phát triển nhận thức của trẻ sau này. Nhiều nghiên cứu đã phát hiện ra rằng trẻ sơ sinh bị thiếu cân có xu hướng có trình độ học vấn thấp hơn, tình trạng sức khỏe kém hơn và có cơ hội việc làm và thu nhập khi trưởng thành thấp hơn so với trẻ có cân nặng bình thường. Do đó, cân nặng của trẻ sơ sinh được coi là thước đo quan trọng về sức khỏe và phúc lợi của trẻ sơ sinh trong đại đa số các phân tích kinh tế học y tế.

*Bài tập này được thiết kế dựa trên nghiên cứu của Almond et al (2005) và bài giảng môn kinh tế lượng ARE 213 của Michael Anderson, University of California at Berkeley.

Trọng lượng trẻ sơ sinh cũng được sử dụng để đánh giá hiệu quả của các chính sách xã hội. Nghiên cứu về lợi ích của các chương trình xã hội quy mô lớn như phúc lợi và bảo hiểm y tế cho người nghèo thường sử dụng cân nặng khi sinh làm chỉ số chính cho phúc lợi trẻ sơ sinh. Mối tương quan giữa LBW và chi phí sinh tại bệnh viện cũng đã được sử dụng để tính toán lợi ích của các can thiệp khuyến khích cai thuốc lá khi mang thai.

Nhìn chung, tình trạng LBW bị chi phối bởi hai yếu tố: thời gian mang thai ngắn (tức là sinh non - preterm) và tốc độ tăng trưởng của thai nhi giảm trong thời gian mang thai (còn được gọi là chậm phát triển trong tử cung - Intrauterine Growth Restriction/IUGR). Trong đó, hành vi hút thuốc trong giai đoạn thai nghén đã được xác định là yếu tố rủi ro có thể ảnh hưởng đáng kể nhất đối với tỷ lệ mắc LBW và tỷ lệ tử vong ở trẻ sơ sinh ở các nước phát triển. Do vậy, đã có những chính sách can thiệp sức khỏe ở Hoa Kỳ nhằm hạn chế hành vi hút thuốc của bà mẹ trong giai đoạn mang thai.

Mô hình kinh tế lượng

Giả sử mô hình đề xuất về mối quan hệ giữa LBW với sức khỏe trẻ sơ sinh là:

$$H_i = \beta_0 + \beta_1 SMOKING_i + \beta_j X_j + u_i \quad (1)$$

Trong đó biến phụ thuộc H là chỉ số đặc trưng cho tình trạng sức khỏe của trẻ sơ sinh; $SMOKING$ là tình trạng hút thuốc của bà mẹ trong khi mang thai; X_j là các biến kiểm soát khác trong mô hình như các đặc tính nhân khẩu học, độ tuổi, học vấn... Bài nghiên cứu này tập trung vào ba chỉ số sức khỏe đặc trưng của trẻ sơ sinh bao gồm:

Bảng 1: Biến Trạng thái Sức khỏe của Trẻ Sơ sinh

Chỉ số sức khỏe	Diễn giải
dbrwt	Trọng lượng khi sinh (gram)
omaps	Chỉ số APGAR* đo sau sinh 1 phút
fmaps	Chỉ số APGAR đo sau sinh 5 phút

* Chỉ số APGAR được xây dựng từ năm tiêu chí (Appearance - màu da, Pulse - nhịp tim, Grimace - phản ứng khi kích thích, Activity - vận động cơ bắp, and Respiration - thở) và nhận giá trị từ 0 đến 10, để mô tả trình trạng sức khỏe của trẻ sơ sinh từ rất xấu đến rất tốt.

Để hoàn thành bài tập này học viên cần đọc file hướng dẫn dữ liệu kèm theo, chọn các biến số phù hợp có thể sử dụng trong mô hình, lọc bộ dữ liệu và xử lý các vấn đề như missing values trước khi ước lượng mô hình. Học viên cần cân nhắc đưa các biến giải thích sau vào mô hình:

Bảng 2: Biến Giải thích

Tên biến	Diễn giải
dmage	tuổi của bà mẹ
dmeduc	trình độ học vấn của bà mẹ
dmar	tình trạng hôn nhân của bà mẹ
monpre	thời điểm bắt đầu nhận chăm sóc mang thai (tháng)
nprevist	số lần đi khám khi mang thai
dgestat	thời gian mang thai (tuần) (gestation)
csex	giới tính của trẻ sơ sinh
dplural	sinh đơn hay sinh đôi, sinh ba...
delmeth5	hình thức sinh (tự nhiên, mổ đẻ/c-section...)
anemia	bà mẹ bị bệnh thiếu máu
cardiac	bà mẹ bị bệnh tim
lung	bà mẹ bị bệnh phổi/hô hấp
diabetes	bà mẹ bị bệnh tiểu đường
herpes	bà mẹ bị nhiễm virus herpes
chyper	bị bệnh căng thẳng thần kinh mãn tính
phyper	bị bệnh căng thẳng thần kinh khi mang thai
preterm	con trước bị đẻ non
tobacco	bà mẹ có hút thuốc lá lúc mang thai
cigar	số điều thuốc lá hút hàng ngày (trung bình)
alcohol	bà mẹ có uống đồ uống có cồn
drink	số lần uống đồ uống có cồn (ly/lần)
wgain	trọng lượng tăng lên trong giai đoạn thai nghén

Câu hỏi

1. Làm sạch bộ dữ liệu bằng các loại bỏ các quan sát có giá trị là missing values. Bộ dữ liệu sau khi làm sạch có 114,610 quan sát. Lập thống kê mô tả và diễn giải.
2. Khi nào thì sự khác biệt của các chỉ số sức khỏe giữa nhóm có hút thuốc và không hút thuốc trong bảng (1) chính là tác động của hành vi hút thuốc? Có bằng chứng nào cho thấy có sự khác biệt mang tính hệ thống giữa nhóm có hút thuốc (nhóm can thiệp) và không hút thuốc không (nhóm đối chứng)?
3. Khi nào thì ước lượng β_1 bằng OLS từ phương trình (1) mang hàm ý quan hệ nhân quả? Giả dụ điều kiện đó thỏa, hãy ước lượng mô hình (1) và diễn giải tác động của hành vi hút thuốc đến ba biến phụ thuộc trong Bảng 1. Trình bày kết quả ước lượng thành bảng biểu, thực hiện các kiểm định cần thiết, và diễn giải ý nghĩa của mô hình.

Để trả lời câu hỏi này, học viên cần đề xuất các biến kiểm soát khác (X_j) đưa vào mô hình ngoài biến *SMOKING*. Lưu ý rằng các biến giải thích đưa vào mô hình (1) phải là các biến ngoại sinh – nghĩa là các biến không bị tác động bởi hành vi

hút thuốc. Biến thời gian mang thai ($dgestat$) và biến dự báo thời gian mang thai ($clingest$) là các biến nội sinh do chúng bị ảnh hưởng trực tiếp bởi hành vi hút thuốc.

4. Giả dụ hành vi hút thuốc là ngẫu nhiên sau khi đã kiểm soát các đặc tính quan sát được ("smoking is randomly assigned conditional on other observable characteristics"). Mô tả cách thức ước lượng mô hình (1) bằng phương pháp ghép cặp (propensity score matching) để tính tác động của hành vi hút thuốc lên sức khỏe của trẻ sơ sinh.
5. Ước lượng tác động của hành vi hút thuốc lên ba biến phụ thuộc bằng phương pháp matching. Giải thích từng bước thực hiện ước lượng và ý nghĩa của ước lượng theo từng thể thức ghép cặp.
6. Dựa trên so sánh giữa các nhóm, ước lượng bằng phương pháp matching cho ra tác động gì (ATT hay ATE)? Thảo luận về hiệu lực ngoại vi của ước lượng bằng matching so với ước lượng bằng OLS hay so sánh thuần túy về mặt trung bình giữa các nhóm.

Học viên có thể sử dụng đồ thị phân phối của các biến giải thích hoặc các kiểm định để so sánh tính tương đồng giữa các nhóm và hiệu lực của phương pháp matching.

Tài liệu tham khảo

Almond, Douglas et al 2005. The cost of low birth weight, *The Quarterly Journal of Economics*, 1031-83.