

ÁP DỤNG PHƯƠNG PHÁP PHÂN BỐ NGẪU NHIÊN TRONG NGHIÊN CỨU CHÍNH SÁCH CÔNG

1. Khái niệm phương pháp phân bố ngẫu nhiên

Hầu hết các can thiệp chính sách được thiết kế và thực hiện trong một môi trường phức tạp và thay đổi, trong đó có nhiều nhân tố có thể tác động đến kết quả đầu ra đối với cả những người tham gia chương trình và những người không tham gia chương trình. Ví dụ hạn hán, động đất, tình trạng suy thoái kinh tế, những thay đổi trong chính quyền, những thay đổi trong các chính sách quốc gia và quốc tế. Do đó các đại biểu cần phải chắc chắn rằng ước lượng của chương trình đạt đến mức độ nào cho dù có vô số nhân tố này.

Các phương pháp phân bố ngẫu nhiên gần giống với việc quay số may rủi để quyết định những người tham gia vào một can thiệp chính sách tại một thời điểm nhất định và những người không tham gia. Các phương pháp phân bố ngẫu nhiên này không chỉ cung cấp cho các cơ quan quản lý chương trình quy tắc công bằng và minh bạch trong việc phân bổ các nguồn lực khan hiếm giữa các nhóm dân số cần được sự trợ giúp một cách bình đẳng mà còn là những phương pháp tốt nhất để đánh giá tác động của can thiệp chính sách. Các phương pháp phân bố ngẫu nhiên thường bắt nguồn từ các quy tắc can thiệp chính sách, dân số của những người tham gia có chủ định tức là tập hợp tất cả các đơn vị mà can thiệp chính sách vụ lớn hơn số lượng người tham gia mà can thiệp chính sách thực sự có thể cung cấp thời điểm xác định. Vì nhiều lý do mà các can thiệp chính sách có thể không có khả năng vươn tới toàn bộ dân số mục tiêu. Chẳng hạn, do những hạn chế ngân sách lên chính phủ không thể cung cấp một chương trình cho tất cả các đơn vị đủ điều kiện từ khi bắt đầu. Ngay cả khi có đủ ngân sách thì những hạn chế về năng lực của các cơ quan cung cấp đôi khi cũng không cho phép một can thiệp chính sách mở rộng cho mọi người một cách đồng thời.

Trong thực tế, hầu hết các can thiệp chính sách gặp những trở ngại ngân sách hoặc năng lực của cơ quan thực hiện lên không thể vươn tới mọi người tham gia mong muốn một cách đồng thời. Trong bối cảnh hiện nay, khi dân số đủ điều kiện tham gia lớn hơn số lượng chỗ trong can thiệp chính sách, thì các cơ quan quản lý chương trình phải xác định cơ chế để phân bổ các dịch vụ của

can thiệp chính sách. Nói cách khác, họ phải quyết định ai sẽ phải tham gia can thiệp chính sách và ai sẽ không tham gia can thiệp chính sách. Can thiệp chính sách có thể được phân bổ trên cơ sở ưu tiên vào đặc điểm quan sát (ví dụ, ưu tiên phụ nữ và trẻ em được phục vụ trước hoặc ưu tiên cộng đồng nghèo nhất được phục vụ trước); hoặc lựa chọn dựa vào các đặc điểm không quan sát được (ví dụ để các cá nhân đăng ký dựa vào động cơ và nhận thức của họ), hoặc thậm chí là quay số may rủi.

Phân bổ ngẫu nhiên được xem là phương pháp thiết thực nhất để ước lượng các phản thực và thường được xem là tiêu chuẩn vàng của đánh giá tác động với một quy mô mẫu đủ lớn, quá trình phân bổ ngẫu nhiên sẽ đảm bảo sự giống hệt nhau trong các đặc điểm quan sát được và không quan sát được giữa hai nhóm can thiệp và nhóm so sánh và do đó xử lý được sai số lựa chọn.

Như vậy, phương pháp phân bổ ngẫu nhiên là phương pháp đánh giá tác động sử dụng phân bổ ngẫu nhiên để lựa chọn những đơn vị vào nhóm can thiệp và nhóm so sánh, và mỗi đơn vị có cơ hội được chọn như nhau.

2. Phân bổ ngẫu nhiên đối tượng can thiệp và phản lực.

2.1 Phân bổ ngẫu nhiên đối tượng can thiệp

Một khi can thiệp chính sách được phân bổ ngẫu nhiên đối với một dân số đủ điều kiện lớn, thì chúng ta có thể tạo ra được một ước lượng hợp lệ về phản thực. Về cơ bản phân ra được một ước lượng hợp lệ về phản thực. Về cơ bản phân bổ ngẫu nhiên đối tượng can thiệp được sử dụng quay số may rủi để quyết định đơn vị nào trong dân số đủ điều kiện được nhận can thiệp chính sách và đơn vị nào thì không. Mọi đơn vị can thiệp đủ điều kiện (ví dụ, cá nhân, hộ gia đình, cộng đồng trường học, bệnh viện, hoặc tổ chức các nhân, hoặc tổ chức khác) có xác suất lựa chọn làm đối tượng can thiệp như nhau.

Sau khi xác định được dân số mục tiêu (ví dụ các hộ gia đình dưới ngưỡng nghèo, hoặc trẻ em dưới 5 tuổi, hoặc các trường học ở các vùng nông thôn), thì phân bổ ngẫu nhiên là một quy tắc phân bổ công bằng bởi vì nó cho phép mọi đơn vị đủ điều kiện có cùng cơ hội nhận được sự can thiệp chính sách và đảm bảo can thiệp chính sách không được phân bổ một cách thức vị thân hoặc không công bằng khác. Khi cầu vượt quá cung của một can thiệp chính sách hiện có, thì phân bổ ngẫu nhiên là một quy tắc giúp cơ quan quản lý chương trình để

dàng giải thích cho người dân hiểu. Khi việc lựa chọn được tiến hành thông qua một quá trình mở và có thể tái tạo, thì quy tắc phân bổ ngẫu nhiên khó bị thao túng và do đó bảo vệ các cơ quan quản lý chương trình khỏi những cáo buộc tiềm năng về sự thiên vị hoặc tham gia nhóm các cá nhân đủ điều kiện, trước hết là vì những ưu điểm của nó đối với quản lý và điều hành chương trình.

2.2 Phân bổ ngẫu nhiên phản thực

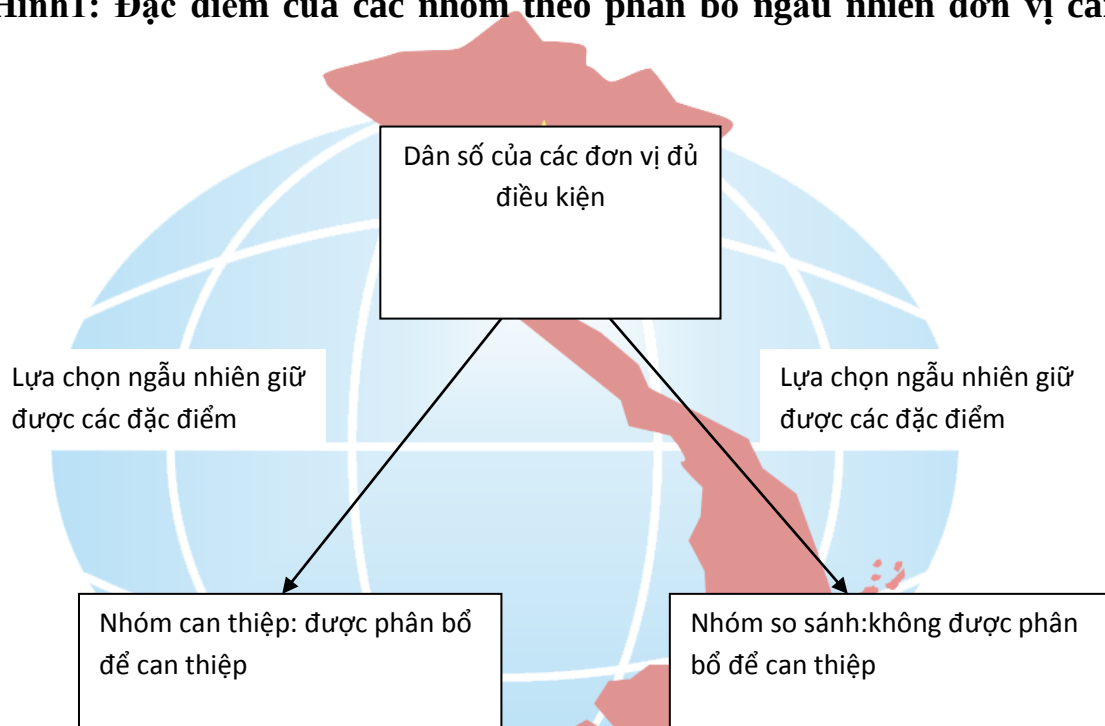
Nhóm so sánh lý tưởng càng giống với nhóm can thiệp ở mọi phương diện càng tốt, ngoại trừ việc nhóm so sánh không tham gia vào can thiệp chính sách đang được đánh giá. Điểm mấu chốt là khi lựa chọn ngẫu nhiên các đơn vị để phân bổ vào nhóm can thiệp và nhóm so sánh, thì bản thân quá trình phân bổ ngẫu nhiên sẽ tạo ra hai nhóm có xác suất thống kê như nhau với điều kiện đơn vị là số lượng đơn vị tham gia tiềm năng cho quá trình lựa chọn ngẫu nhiên là đủ lớn. Cụ thể với một số lượng quan sát đủ lớn, quá trình phân bổ ngẫu nhiên sẽ tạo ra các nhóm, về thống kê, có các đặc điểm trung bình tương đương đối với tất cả các đặc điểm của chúng. Các đặc điểm trung bình tương đương này cũng giống với đặc điểm trung bình của dân số mục tiêu

Hình 1 giải thích lý do phân bổ ngẫu nhiên tạo ra một nhóm so sánh, về thống kê, tương đương với nhóm can thiệp giả sử dân số của đơn vị đủ điều kiện (những người tham gia tiềm năng) gồm 500 người, trong đó một nửa được lựa chọn ngẫu nhiên và phân bổ vào nhóm can thiệp và nửa còn lại phân bổ vào nhóm so sánh. Ví dụ, chúng ta có thể viết tên của 500 người lên từng mảnh giấy và trộn lẫn chúng lại với nhau, sau đó yêu cầu một người bất kỳ bị bịt mắt lấy ra 250 mảnh giấy. Nếu xác định được 250 cái tên đầu tiên sẽ tạo ra nhóm can thiệp được phân bổ ngẫu nhiên (250 cái tên đầu tiên được lấy ra), và một nhóm so sánh được phân bổ ngẫu nhiên (250 cái tên còn lại).

Giả sử trong 500 người ban đầu có 60% là phụ nữ. Bởi vì các mảnh giấy được lựa chọn ngẫu nhiên, nên 250 cái tên được lấy ra có xấp xỉ 60% là phụ nữ. Nói chung, nếu dân số của các đơn vị đủ điều kiện là đủ lớn, thì mọi đặc điểm của dân số sẽ được phân bổ đều cho các nhóm can thiệp và nhóm so sánh, thì về logic các đặc điểm khó quan sát hơn (các biến không quan sát được) như động cơ, sở thích, hoặc đặc điểm tính cách khó đo lường khác, sẽ cũng được phân bổ đều cho cả nhóm can thiệp và nhóm so sánh. Do đó, nhóm can thiệp và nhóm so sánh được tạo ra thông qua phân bổ ngẫu nhiên sẽ giống hệt nhau không chỉ về

đặc điểm quan sát được. Ví dụ, chúng ta không thể quan sát hoặc đo lường người “tốt” là như thế nào, nhưng chúng ta biết rằng nếu 40% số người trong dân số của đơn vị đủ điều kiện là tốt, thì có xấp xỉ 40% số người trong nhóm can thiệp sẽ là tốt và 40% số người trong nhóm so sánh là tốt. Phân bổ ngẫu nhiên sẽ giúp đảm bảo, về trung bình, nhóm can thiệp và nhóm so sánh là tương đồng ở mọi phương diện, trong cả các đặc điểm quan sát và không quan sát được.

Hình1: Đặc điểm của các nhóm theo phân bổ ngẫu nhiên đơn vị can thiệp

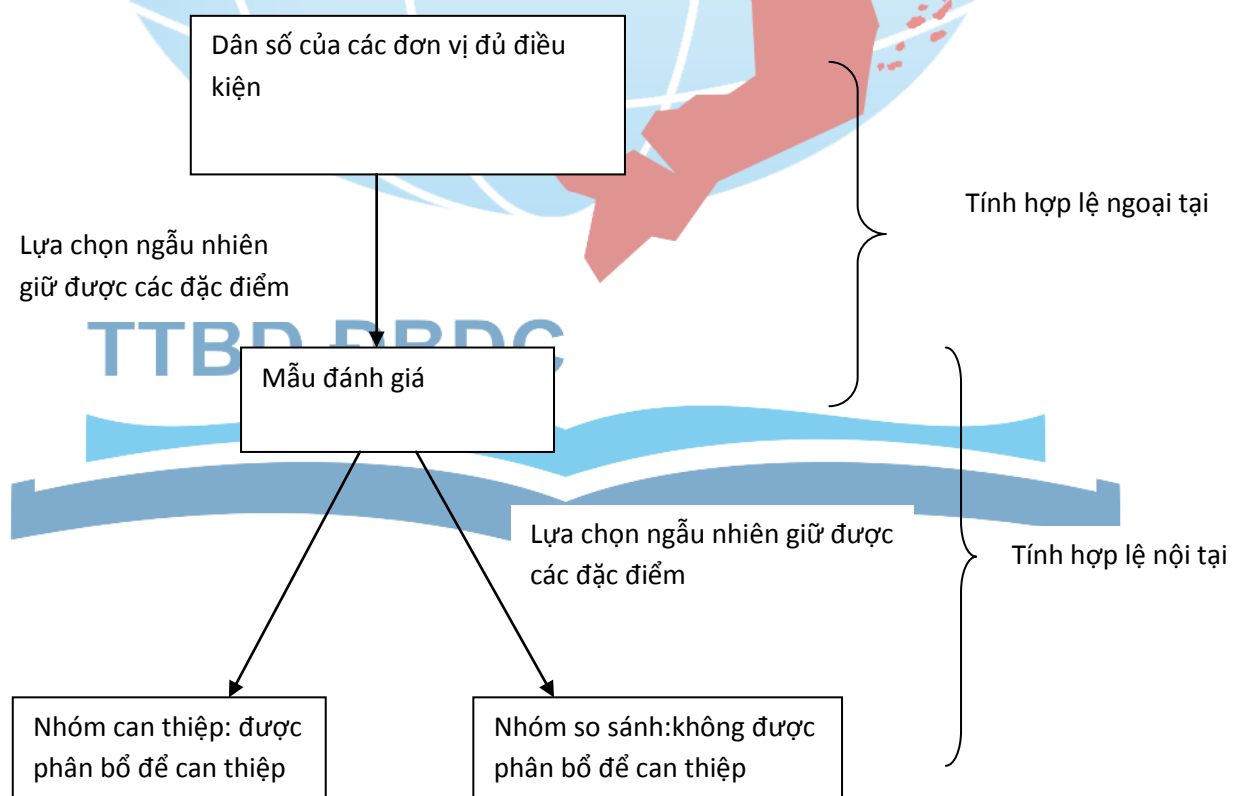


Nguồn: tham khảo từ Impac Evaluation in Practice, The World Bank, p52

Khi một đánh giá sử dụng phân bổ ngẫu nhiên vào nhóm can thiệp và nhóm so sánh, thì về lý thuyết quá trình này sẽ tạo ra hai nhóm tương đương nhau. Với dữ liệu cơ sở về mẫu đánh giá, chúng ta có thể kiểm tra giả định này bằng thực nghiệm và chứng minh trong thực tế không có những khác biệt mang tính hệ thống trong các đặc điểm quan sát giữa các nhóm can thiệp và nhóm so sánh trước khi can thiệp chính sách bắt đầu. Do đó, sau khi can thiệp chính sách khởi động, nếu chúng ta quan sát những khác biệt trong những kết quả đầu ra giữa các nhóm can thiệp và nhóm so sánh thì những khác biệt đó chỉ có thể giải thích là do can thiệp chính sách, bởi vì khi xây dựng hai nhóm đã giống hệt nhau ở dữ liệu cơ sở và đã bộc lộ trước các nhân tố môi trường bên ngoài giống nhau qua thời gian. Từ góc độ này, nhóm so sánh kiểm soát tất cả các nhân tố có thể tác động đến kết quả đầu ra mong muốn. Chúng ta có thể hoàn toàn tin tưởng chênh

lệch giữa các kết quả đầu ra đối với nhóm can thiệp (kết quả đầu ra trung bình của nhóm can thiệp được phân bổ ngẫu nhiên) chính là tác động thực của can thiệp chính sách, bởi vì khi xây dựng mẫu chúng ta loại bỏ tất cả các nhân tố quan sát được và không quan sát được mà có thể tác động đến kết quả đầu ra. Hình 1 giả định tất cả các đơn vị trong dân số đủ điều kiện sẽ được phân bổ hoặc vào nhóm can thiệp hoặc vào nhóm so sánh. Tuy nhiên, trong một số trường hợp, chúng ta không nhất thiết phải đánh giá chúng. Ví dụ, nếu dân số của các đơn vị đủ điều kiện gồm hai triệu người mẹ, và chúng ta muốn đánh giá tác động của chương trình thưởng tiền mặt lên xác suất của việc đưa con của họ đi tiêm chủng, thì chúng ta có thể chỉ cần lấy mẫu đại diện là 1.500 người mẹ là đủ và phân bổ 1.500 người mẹ này vào nhóm can thiệp hoặc nhóm so sánh như hình minh họa hình 2. Cũng theo cách giải thích logic ở trên, việc lấy một mẫu ngẫu nhiên từ dân số của các đơn vị đủ điều kiện để thành lập mẫu đánh giá giữ được các đặc điểm của dân số của các đơn vị đủ điều kiện. Việc lựa chọn ngẫu nhiên nhóm can thiệp và nhóm so sánh từ mẫu đánh giá giữ lại được các đặc điểm.

Hình 2. chọn ngẫu nhiên và phân bổ ngẫu nhiên nhóm can thiệp



Lựa chọn ngẫu nhiên
giữ được các đặc điểm

Nguồn: tham khảo từ Impac Evaluation in Practice, The World Bank (2011), p.54.

3. Tính hợp lệ nội tại và hợp lệ ngoại tại

Việc phân bổ ngẫu nhiên nhóm can thiệp sẽ bảo đảm cả tính hợp lệ ngoại tại và nội tại của đánh giá tác động, với điều kiện là mẫu đánh giá đủ lớn.

- Tính hợp lệ nội tại: Có nghĩa là tác động ước lượng của can thiệp chính sách là tác động còn lại sau khi đã loại bỏ tác động các nhân tố gây xáo trộn tiềm năng khác, hoặc có nghĩa là nhóm so sánh thể hiện phản thực hợp lệ, do đó, chúng ta ước lượng được tác động thực sự của can thiệp chính sách. Về thống kê, phân bổ ngẫu nhiên tạo ra nhóm so sánh độc lập cùng tập hợp các nhân tố bên ngoài qua thời gian. Do đó, nếu có những khác biệt trong các kết quả đầu ra xuất hiện giữa nhóm can thiệp này và nhóm so sánh, thì chúng chỉ có thể là do nhóm can thiệp tham gia can thiệp chính sách. Nói cách khác, tính hợp lệ nội tại của một đánh giá tác động được bảo đảm thông qua quá trình phân bổ ngẫu nhiên nhóm can thiệp. Một đánh giá tác động sẽ bảo đảm tính hợp lệ nội tại nếu chúng ta sử dụng một nhóm so sánh hợp lệ.

- Tính hợp lệ ngoại tại: Có nghĩa là tác động ước lượng được trong mẫu đánh giá có thể suy rộng ra cho dân số của tất cả các đơn vị đủ điều kiện. Để có được tính hợp lệ ngoại tại, thì mẫu đánh giá phải là đại diện cho dân số của các đơn vị đủ điều kiện; trong thực tế, mẫu đánh giá phải được, lựa chọn từ dân số bằng việc sử dụng một trong cách thức lựa chọn ngẫu nhiên.

Có hai loại mẫu ngẫu nhiên khác nhau: Một mẫu ngẫu nhiên cho mục đích chọn mẫu (đối với tính hợp lệ ngoại tại) và một mẫu ngẫu nhiên cho đánh giá tác động (đối với tính hợp lệ nội tại). Một đánh giá tác động có thể tạo ra các ước lượng về tác động hợp lệ nội tại thông qua phân bổ ngẫu nhiên nhóm can thiệp; tuy nhiên, nếu đánh giá được tiến hành trên mẫu dân số không ngẫu nhiên, thì các tác động ước lượng có thể không có khả năng suy rộng ra cho dân số của các đơn vị đủ điều kiện. Ngược lại, nếu đánh giá tác động sử dụng mẫu ngẫu nhiên của dân số của các đơn vị đủ điều kiện, nhưng nhóm can thiệp không được phân bổ ngẫu nhiên thì mẫu sẽ là đại diện nhưng nhóm so sánh không thể không hợp lệ. Một đánh giá là hợp lệ ngoại tại nếu mẫu đánh giá hoàn toàn đại diện cho dân số của các đơn vị đủ điều kiện. Do đó, những kết quả đánh giá có thể suy rộng cho dân số của các đơn vị đủ điều kiện.

4. Trường hợp sử dụng phương pháp phân bổ ngẫu nhiên

Trong thực tế, phương pháp phân bổ ngẫu nhiên cần được cân nhắc trong những trường hợp nhu cầu vượt quá cung của chính sách, tức là khi số lượng người tham gia, có tiềm năng lớn hơn khả năng đáp ứng của can thiệp chính sách tại một thời điểm nhất định và can thiệp chính sách cần được phân đoạn. Một số trường hợp cho dù khả năng của can thiệp chính sách không bị hạn chế, nhưng chúng ta cũng cần sử dụng phân bổ ngẫu nhiên làm công cụ đánh giá. Ví dụ, Chính phủ muốn sử dụng phân bổ ngẫu nhiên để kiểm tra các chính sách mới hoặc các chương trình có khả năng là tốn kém và không dự đoán trước được các hậu quả. Trong trường hợp này, phương pháp phân bổ ngẫu nhiên là cần thiết trong giai đoạn đánh giá thí điểm để kiểm tra nghiêm ngặt các tác động của chính sách trước khi mở rộng nó. Thông thường có hai loại trường hợp có thể sử dụng phương pháp phân bổ ngẫu nhiên để đánh giá các tác động.

- Khi dân số đủ điều kiện lớn hơn khả năng đáp ứng của chương trình: Khi cầu đối với một can thiệp chính sách vượt quá cung của nó thì người đánh giá có thể sử dụng phân bổ ngẫu nhiên để lựa chọn nhóm can thiệp trong phạm vi dân số đủ điều kiện. Trong trường hợp này, mọi đơn vị trong dân số có cơ hội được lựa chọn tham gia vào can thiệp chính sách như nhau. Nhóm có được may mắn sẽ là nhóm can thiệp và phần dân số còn lại không được nhận can thiệp chính sách là nhóm so sánh. Với điều kiện: Hạn chế về nguồn lực không cho phép mở rộng can thiệp chính sách cho toàn bộ dân số, nhóm so sánh có thể được duy trì để đo lường các tác động ngắn hạn, trung hạn và dài hạn. Trong trường hợp này, việc đánh giá tác động không gặp phải vấn đề đạo đức nghiêm trọng do phải duy trì nhóm so sánh vô hạn định, bởi vì một bộ phận dân số sẽ phải ở ngoài can thiệp chính sách.

Ví dụ, giả sử bộ giáo dục và Đào tạo muốn cung cấp thiết bị dạy học cho các trường công lập trên toàn quốc, nhưng Bộ tài chính cho rằng ngân sách chỉ đủ cấp cho 1/3 số trường. Nếu Bộ giáo dục và Đào tạo muốn mỗi trường công có cơ hội nhận được thiết bị dạy học như nhau, thì bốc thăm may rủi, theo đó mỗi trường có cơ hội lựa chọn là 1/2. Các trường may mắn nhận được thiết bị dạy học mới và tạo thành nhóm can thiệp, và 2/3 số trường còn lại không được cung cấp thiết bị dạy học được xem là nhóm so sánh. Như vậy, một nhóm các

trường công không được chương trình tài trợ thiết bị dạy học và nó được sử dụng làm nhóm so sánh để đo lường phản thực.

- Khi chương trình cần phân đoạn đều đặn cho đến khi bao phủ toàn bộ dân số đủ điều kiện: Khi một can thiệp chính sách được phân thành các giai đoạn, thì việc lựa chọn ngẫu nhiên thứ tự những đơn vị nhận được can thiệp chính sách cho phép từng đơn vị đủ điều kiện có cùng cơ hội là nhóm can thiệp trong giai đoạn đầu hoặc cùng một giai đoạn của can thiệp chính sách. Với điều kiện: Nhóm “cuối cùng” không được phân đoạn vào can thiệp chính sách và được xem là nhóm so sánh hợp lệ, và thông qua đó chúng ta ước lượng được phản thực đối với các nhóm đã được phân vào các giai đoạn.

Ví dụ, giả sử Bộ giáo dục và Đào tạo muốn đào tạo nâng cao tiếng Anh cho tất cả các giáo viên dạy tiếng Anh ở các trường công lập trong toàn quốc thì phải cần thời gian 3 năm. Để tiến hành đánh giá tác động, Bộ giáo dục và Đào tạo có thể chọn ngẫu nhiên 1/3 số giáo viên được đào tạo trong năm thứ nhất, 1/3 trong năm thứ hai và 1/3 trong năm thứ ba. Để đánh giá tác động của chương trình đào tạo sau một năm thực hiện, nhóm giáo viên được đào tạo trong năm thứ nhất sẽ tạo thành nhóm can thiệp và nhóm giáo viên được phân bổ ngẫu nhiên vào đào tạo trong năm thứ ba sẽ là nhóm so sánh bởi vì họ đã không nhận được chương trình đào tạo trong năm thứ nhất và thứ hai.

5. Các cấp độ phân bổ ngẫu nhiên

Phân bổ ngẫu nhiên có thể được thực hiện ở cấp độ cá nhân, hộ gia đình, cộng đồng hoặc vùng. Nói chung, cấp độ phân bổ ngẫu nhiên các đơn vị vào nhóm can thiệp và nhóm so sánh sẽ bị tác động lớn bởi nơi và cách thức thực hiện can thiệp chính sách. Ví dụ, nếu một chương trình giáo dục đang được thực hiện ở cấp trường tiểu học và sau đó phân bổ ngẫu nhiên một số trường vào nhóm can thiệp và một số trường khác vào nhóm so sánh.

Khi cấp độ phân bổ ngẫu nhiên là cao hơn, ví dụ, ở vị trí cấp vùng hoặc tỉnh trong một quốc gia thì rất khó khăn để thực hiện đánh giá tác động bởi vì số lượng vùng và tỉnh hầu hết các quốc gia là không đủ lớn để thu được nhóm can thiệp và nhóm so sánh cân bằng. Ví dụ, nếu một quốc gia chỉ có 10 tỉnh, nên chỉ cho 5 tỉnh can thiệp và 5 tỉnh so sánh, thì số lượng này không đủ để đảm bảo rằng các đặc điểm của nhóm can thiệp và nhóm so sánh cân bằng nhau. Nhưng ở

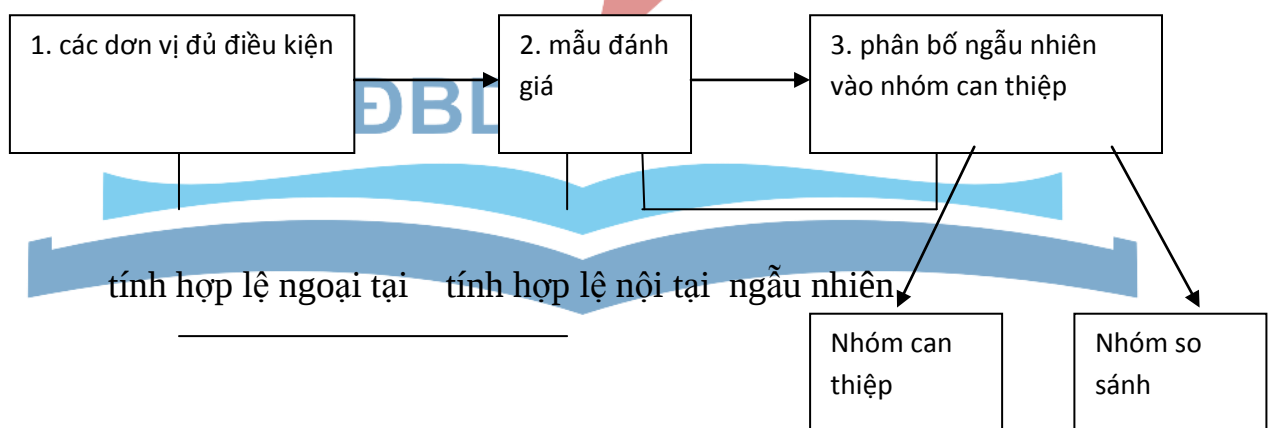
các cấp độ phân bổ ngẫu nhiên thấp hơn ví dụ xuống đến cấp cá nhân, hoặc cấp hộ gia đình, có cơ hộ lan tỏa và bị làm bẩn mẫu tăng lên. Ví dụ, một chương trình cung cấp thuốc diệt chuột cho các hộ gia đình và một hộ gia đình trong nhóm can thiệp ở rất gần một hộ trong nhóm so sánh thì hộ trong nhóm so sánh có thể được tác động tích cực bởi sự lan tỏa từ sự can thiệp, đã cung cấp cho nhóm can thiệp bởi vì hộ có chuột từ những nhà hàng xóm sẽ giảm đi. Các can thiệp và so sánh cần có vị trí đủ xa nhau để tránh sự lan tỏa như thế. Khi khoảng cách giữa các hộ gia đình tăng lên, thì chi phí tốn kém hơn so với cả việc thực hiện chương trình và cho cả việc tiến hành các khảo sát. Theo quy tắc ngón tay cái, nếu sự lan tỏa có thể kiểm soát một cách hợp lý, thì tốt nhất là thực hiện phân bổ ngẫu nhiên nhóm can thiệp ở cấp độ thực hiện chương trình thấp nhất; điều đó sẽ đảm bảo số lượng đơn vị trong cả nhóm can thiệp, và nhóm so sánh càng lớn càng tốt.

6. Tiến hành đánh giá tác động theo phương pháp phân bổ ngẫu nhiên

6.1. Tiến hành phân bổ ngẫu nhiên nhóm can thiệp

Quy trình phân bổ ngẫu nhiên nhóm can thiệp gồm ba bước như được mô tả ở hình 3.

Hình 3: quy trình phân bổ ngẫu nhiên để tạo thành nhóm can thiệp



Nguồn: tham khảo từ Impac Evaluation in Practice, The Word Bank (2011), tr 57

- Bước 1: Xác định các đơn vị đủ điều kiện tham gia can thiệp chính sách. Tùy thuộc vào can thiệp chính sách cụ thể, một đơn vị có thể là một gia đình

một tổ chức, một cộng đồng... Dân số của các đơn vị đủ điều kiện sẽ là dân số mà chúng ta muốn biết tác động của can thiệp chính sách. Ví dụ, nếu chúng ta đang thực hiện một chương trình tiếng anh nâng cao cho các giáo viên tiếng anh trong các trường tiểu học, trong các trường phổ thông cơ sở ở các vùng nông thôn, hoặc các giáo viên tiếng anh trong tất cả các trường ở vùng đô thị sẽ không phụ thuộc vào dân số của các đơn vị đủ điều kiện. Sau khi xác định được dân số của các đơn vị đủ điều kiện chúng ta cần so sánh quy mô của một nhóm với số lượng quan sát cần cho đánh giá. Số lượng này được xác định thông qua tính toán mẫu tối thiểu và dựa vào loại câu hỏi chính sách. Nếu dân số đủ điều kiện là nhỏ, thì tất cả các đơn vị đủ điều kiện có thể được bao gồm trong đánh giá. Nếu số đơn vị đủ điều kiện lớn hơn số đơn vị cần thiết để đánh giá, thì chuyển sang bước 2.

- Bước 2: Lựa chọn mẫu đánh giá. Trong bước này, chúng ta chọn mẫu các đơn vị từ dân số vào mẫu đánh giá. Bước này được tiến hành chủ yếu nhằm hạn chế chi phí thu thập dữ liệu. Nếu dữ liệu từ các hệ thống giám sát có thể sử dụng được cho đánh giá và các hệ thống đó bao phủ dân số của các đơn vị đủ điều kiện, thì chúng ta không cần rút ra một mẫu đánh giá riêng biệt. Tuy nhiên, khi tiến hành một đánh giá bao gồm dân số mục tiêu của các đơn vị đủ điều kiện với hàng chục nghìn giáo viên tiếng anh trung học phổ thông ở các vùng nông thôn trong toàn quốc, chúng ta cần thu thập thông tin chi tiết về kiến thức tiếng anh của họ, thì việc phỏng vấn tất cả những giáo viên tiếng anh này có thể không khả thi trong thực tế, nên chúng ta có thể chỉ cần một mẫu 1.500 giáo viên tiếng anh được phân bổ qua 100 trường trung học phổ thông ở các vùng nông thôn. Với điều kiện: Mẫu các trường trung học phổ thông và giáo viên tiếng anh của các trường trung học phổ thông đủ điều kiện, mọi kết quả tìm được trong đánh giá có thể suy rộng cho tất cả các giáo viên tiếng anh và trường trung học phổ thông ở các vùng nông thôn còn lại trong toàn quốc. Việc thu thập dữ liệu về mẫu 1.500 giáo viên tiếng anh sẽ thấp hơn việc thu thập dữ liệu về tất cả các trường trung học phổ thông ở các vùng trong toàn quốc.

- Bước 3: phân bổ ngẫu nhiên vào nhóm can thiệp. trong bước này, chúng ta xây dựng nhóm can thiệp và nhóm so sánh các số đơn vị trong mẫu đánh giá. Để thực hiện bước này, chúng ta phải quyết định quy tắc về cách thức phân bổ những đơn vị tham gia dựa vào con số ngẫu nhiên. Ví dụ, nếu chúng ta cần phân bổ 50 trong 100 đơn vị từ mẫu đánh giá vào nhóm can thiệp, thì chúng ta có thể

quyết định 50 đơn vị có con số ngẫu nhiên cao nhất vào nhóm can thiệp các con số còn lại vào nhóm so sánh. Sau đó chúng ta phân bổ một con số ngẫu nhiên cho từng đơn vị quan sát trong mẫu đánh giá, sử dụng bảng tính hoặc phần mềm thống kê chuyên sâu, sử dụng quy tắc lựa chọn trước đó để hình thành nhóm can thiệp và nhóm so sánh. Lưu ý chúng ta phải quyết định về quy tắc phân bổ trước khi chạy phần mềm gán cho các đơn vị những con số ngẫu nhiên; mặt khác nếu chúng ta quyết định về quy tắc phân bổ dựa vào những con số ngẫu nhiên mà chúng ta nhìn thấy thì sẽ làm mất giá trị của phân bổ ngẫu nhiên. Trong những trường hợp đòi hỏi tiến hành phân bổ ngẫu nhiên công khai trước người dân, thì chúng ta có thể sử dụng một số kỹ thuật thủ công để phân bổ ngẫu nhiên. Dưới đây là một số kỹ thuật thủ công để phân bổ ngẫu nhiên các đơn vị là cá nhân:

- Nếu cần phân bổ 50% các cá nhân vào nhóm can thiệp và 50% vào nhóm so sánh, thì đối với từng người chúng ta tung một lần đồng xu. Trước khi làm việc đó, chúng ta phải quyết định mặt sấp hay mặt ngửa của đồng xu sẽ phân bổ một người vào nhóm can thiệp.

- Nếu cần phân bổ 1/3 mẫu đánh giá vào nhóm can thiệp, thì đối với từng người chúng ta có thể gieo một lần con xúc xắc. Trước khi làm việc đó, chúng ta phải quyết định phân bổ. Chẳng hạn, mặt trên của con xúc xắc chỉ dừng chỉ 1 hoặc 2 chấm thì người đó được phân bổ vào nhóm can thiệp, và chỉ 3, 4, 5 hoặc 6 chấm thì người đó được phân bổ vào nhóm so sánh. Thứ hai, chúng ta gieo con xúc xắc một lần đối với một người trong mẫu đánh giá và căn cứ vào con số có được để có thể phân bổ họ vào nhóm can thiệp hoặc nhóm so sánh.

- Viết tên của từng người lên một mẫu giấy có hình thù và kích thước như nhau. Gập mẫu giấy để không thể nhìn thấy tên, trộn kỹ chúng trong một chiếc mũ hoặc cái hộp. Chúng ta quyết định quy tắc phân bổ trước khi bắt đầu lấy từng mảnh giấy ra, tức là quyết định bao nhiêu mẫu giấy sẽ được lấy ra và tên của người được lấy sẽ được phân bổ vào nhóm can thiệp. Khi quy tắc đã rõ ràng, chọn một người khách quan trong nhóm lấy ra số lượng, mẫu giấy cần để xác định tên những người tham gia vào nhóm can thiệp.

Cho dù sử dụng tung đồng xu, gieo xúc xắc, hoặc con số ngẫu nhiên do máy tính tạo ra,... thì điều quan trọng là chúng minh được, quá trình đó có bảo đảm minh bạch. Nghĩa là: Thứ nhất, quy tắc phân bổ được quyết định trước và thông tin cho mọi thành viên chứng kiến. Thứ hai, phải bám vào quy tắc khi lấy ra con số ngẫu nhiên. Thứ ba, bảo đảm quá trình lấy ra hoàn toàn ngẫu nhiên.

Trong các trường hợp tung đồng xu và gieo xúc xắc, có thể lặp lại quá trình; phân bổ các con số ngẫu nhiên bằng máy tính đòi hỏi cung cấp lo ga của các phép tính, với mục đích để quá trình có thể tái tạo lại bởi những người kiểm tra.

6.2. Ước lượng tác động theo phương pháp phân bổ ngẫu nhiên

Sau khi có được mẫu đánh giá ngẫu nhiên và nhóm can thiệp được phân bổ theo các thức ngẫu nhiên thù chúng ta hoàn toàn dễ dàng ước lượng được tác động của can thiệp chính sách. Sau khi can thiệp chính sách được thực hiện trong một khoảng thời gian nào đó, các kết quả đầu ra đối với cả nhóm can thiệp và nhóm so sánh cần được đo lường. Tác động của can thiệp chính sách đơn giản là sự khác biệt giữa các kết quả đầu ra bình quân (Y) đối với nhóm can thiệp và kết quả đầu ra (Y) đối với nhóm so sánh. Ví dụ, kết quả đầu ra trung bình đối với nhóm so sánh là 75, thì tác động chương trình là 25.

